



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 98597

TO: Phillip Gabel
Location: 8b03 / 9e12
Monday, July 14, 2003
Art Unit: 1644
Phone: 308-3997
Serial Number: 09 / 825580

From: Jan Delaval
Location: Biotech-Chem Library
CM1-1E07
Phone: 308-4498

jan.delaval@uspto.gov

Search Notes

09 / 825580

Jan Delaval
Reference Librarian
Biotechnology & Chemical Library
CM1 1E07 - 703-308-4498
jan.delaval@uspto.gov

282 285 288

68543

Delaval, Jan

From: Gambel, Phillip
Sent: Sunday, July 13, 2003 9:36 AM
To: Delaval, Jan
Subject: 09 / 825580 eppihimer amd

jan

please perform a sequence and a sequence interference search for
09 / 825,580 (eppihimer amd)

SEQ ID NO: 2

thanx

phillip gambel
art unit 1644
308-3997

1644 mailbox 9e12

Jan Delaval
Reference Librarian
Biotechnology & Chemical Library
CM1 1E07 - 703-308-4498
jan.delaval@uspto.gov

THIS PAGE BLANK (USPTO)

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 An Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

Jan Delaval
 Reference Librarian
 Biotechnology & Chemical Library
 CM1 1E07 - 703-308-4498
 jan.delaval@uspto.gov

STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher: <u>Ja</u>	NA Sequence (#) _____	STN _____
Searcher Phone #: <u>4498</u>	AA Sequence (#) <input checked="" type="checkbox"/>	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel Orbit _____
Date searched: <u>7/14/03</u>	Bibliographic _____	Dr Link _____
Date indexed: <u>7/14/03</u>	Litigation _____	Lexis Nexis _____
Searcher Prep & Review Time: _____	Fulltext: _____	Sequence Systems <input checked="" type="checkbox"/>
Client Prep Time: <u>10</u>	Patent Family _____	WWW Internet _____
Indexing Time: <u>10</u>	Other _____	Other Specialty _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: July 14, 2003, 06:10:58 ; Search time 41 Seconds
(without alignments)
942.586 Million cell updates/sec

Title: US-09-825-580-2
Perfect score: 2030
Sequence: 1 MFLQLLLLLLLGPGNSLQL.....TPEPREDDDLTLHSLFLP 402

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2015	99.3	412	A57468	P-selectin glycoprotein
2	288.5	14.2	1	GSFF3	salivary glue prot
3	283	13.9	867	T45463	membrane glycoprot
4	271.5	13.4	866	T45462	membrane glycoprot
5	256	12.6	750	T42614	probable envelope
6	250.5	12.3	797	VBEX1	glycoprotein X pre
7	226.5	11.2	801	T29018	hypothetical prote
8	210.5	10.4	825	T29634	hypothetical prote
9	205.5	10.1	1151	T18535	high molecular mas
10	204.5	10.1	770	T22808	hypothetical prote
11	198	9.8	3191	T22945	hypothetical prote
12	194.5	9.6	507	T44768	antifreeze glycope
13	194	9.6	1161	S57180	probable membrane
14	193.5	9.5	796	T21460	hypothetical prote
15	193	9.5	658	T41309	hypothetical threo
16	190	9.4	2761	T21064	hypothetical prote
17	189.5	9.3	3570	T45025	mucin MUC5B, trach
18	188.5	9.3	1229	T25697	hypothetical prote
19	188.5	9.3	1829	T24583	hypothetical prote
20	187	9.2	1367	S48478	glucan 1,4-alpha-g
21	186.5	9.2	3020	A43922	mucin 2 precursor,
22	186	9.2	402	E86185	hypothetical prote
23	184	9.1	662	A45155	mucin FIM-C.1 - Af
24	183	9.0	1777	T34369	hypothetical prote
25	182.5	9.0	851	T22696	hypothetical prote
26	182.5	9.0	1832	T31113	mucin-like glycopr
27	181	8.9	235	PC2022	mucin like protein
28	180.5	8.9	400	A28172	spasmodic precur
29	180.5	8.9	839	F75518	hypothetical prote

30	180	8.9	3507	2	T34513	hypothetical prote
31	179.5	8.8	328	2	S01359	salivary glue prot
32	178	8.8	292	2	S24169	mucin - rat
33	176.5	8.7	279	2	S53363	mucin 5AC (clone J
34	176.5	8.7	699	2	C43674	US4 protein - huma
35	175	8.6	846	2	T21700	hypothetical prote
36	174.5	8.6	232	2	A60095	larval glue protei
37	173.5	8.5	294	2	A37232	mucin, tracheal (A
38	173.5	8.5	860	2	JC4566	chitinase (EC 3.2.
39	173	8.5	216	2	I51920	mucin - rhesus mac
40	173	8.5	2225	2	T26063	hypothetical prote
41	171.5	8.4	792	2	S70305	hypothetical prote
42	171	8.4	648	2	PC4395	mucin 3 - human (E
43	171	8.4	5376	2	T42215	zonadhesin - mouse
44	170.5	8.4	354	2	T46740	microfilarial shea
45	170	8.4	246	2	PC4397	mucin 3 T10 - huma

ALIGNMENTS

RESULT 1

A57468
P-selectin glycoprotein ligand PSGL-1 precursor, long splice form - human
N:contains: P-selectin glycoprotein ligand PSGL-1, long splice form; P-selectin glycopro
C:Species: Homo sapiens (man)
C>Date: 08-Dec-1995 #sequence revision 08-Dec-1995 #text_change 08-Oct-1999
C:Accession: A57468; A49267; A54711
R:Veldman, G.M.; Bean, K.M.; Cumming, D.A.; Eddy, R.L.; Sait, S.N.J.; Shows, T.B.
J. Biol. Chem. 270, 16470-16475, 1995
A>Title: Genomic organization and chromosomal localization of the gene encoding human P-
A:Reference number: A57468; MUID:95332364; PMID:7541799
A:Accession: A57468
A:Molecule type: DNA
A:Residues: 1-412 <VEL>
R:Cross-references: GB:U95955
R:Sako, D.; Chang, X.J.; Barone, K.M.; Vachino, G.; White, H.M.; Shaw, G.; Veldman, G.M.
Cell 75, 1179-1186, 1993
A>Title: Expression cloning of a functional glycoprotein ligand for P-selectin.
A:Reference number: A49267; MUID:94084793; PMID:7505206
A:Accession: A49267
A>Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-134,145-412 <SAK>
A:Cross-references: GB:U02297; NID:9435416; PIDN:AAG50061.1; PID:9435417
R:Moore, K.L.; Eaton, S.F.; Lyons, D.E.; Lichenstein, H.S.; Cummings, R.D.; McEver, R.P.
J. Biol. Chem. 269, 23318-23327, 1994
A>Title: The P-selectin glycoprotein ligand from human neutrophils displays sialylated,
A:Reference number: A54711; MUID:94365038; PMID:7521878
A:Accession: A54711
A:Molecule type: protein
A:Residues: 350-355;390-391,'K',393-396 <MOO>
C:Genetics: GDB:SELPLG
A:Gene: GDB:SELPLG
A:Cross-references: GDB:273698; OMIM:600738
A:Map position: 12q24-12q24
A:Introns: #status absent
C:Keywords: disulfide bond; glycoprotein; polymorphism; tandem repeat; transmembrane pro
F:1-412/Product: P-selectin glycoprotein ligand PSGL-1 precursor, long splice form #stat
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-319/Domain: extracellular #status predicted <EXT>
F:117-261/Region: 10-residue repeats (A-T/M-E-A-Q-T-X-P/L-A/T)
F:320-343/Domain: transmembrane #status predicted <TM>

Query Match 99.3%; Score 2015; DB 2; Length 412;
Best Local Similarity 97.6%; Pred. No. 2.1e-119;
Matches 402; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 MFLQLLLLLLLGPGNSLQLWDTWADEAKALGPLLARDRQATEYELVDYDFLPETEPP 60
|||||
DB 1 MFLQLLLLLLLGPGNSLQLWDTWADEAKALGPLLARDRQATEYELVDYDFLPETEPP 60

QY 344 -----VRNYSPTMVCISLLPDCGGGFSATANGLSKA 377
 Db 626 GPGNYTFRATMTDDKKVYTYANVYIQEYSGTIESESS-----TSAVASSTSTP 677
 QY 378 KSPGLT-----PPREDREGDDLIT 396
 Db 678 STPSTLSTSTVTEPSTRSSDSTT 702

Query Match 9.6%; Score 194; DB 2; Length 1161;
 Best Local Similarity 26.6%; Pred. No. 0.00012;
 Matches 73; Conservative 30; Mismatches 125; Indels 46; Gaps 11;

RESULT 15
 T41309
 hypothetical threonine/asparagine-rich protein - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 08-Dec-2000
 C:Accession: T41309; T40875
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Wedler, H.; Wambutt, R.; Wedler, E.
 submitted to the EMBL Data Library, March 1998
 A:Reference number: Z21986
 A:Accession: T41309
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-658 <W001>
 A:Cross-references: EMBL:AL022245; NID:G4584240; PIDN:CAA18304.1; PID:G2995364; GSPDB:G

A:Experimental source: strain 972h.; cosmid c320
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Murphy, L.; Harris, D.
 submitted to the EMBL Data Library, September 1998
 A:Reference number: Z21954
 A:Accession: T40875
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 468-658 <W002>
 A:Cross-references: EMBL:AL031764; NID:G3668147; PIDN:CAA21105.1; PID:G3668148; GSPDB:G

A:Experimental source: strain 972h.; cosmid c1235
 C:Genetics:
 A:Gene: SPDB:SPCC320.02c; SPDB:SPCC1235.01
 A:Map position: 3
 Query Match 9.5%; Score 193; DB 2; Length 658;
 Best Local Similarity 29.6%; Pred. No. 7.5e-05;
 Matches 87; Conservative 33; Mismatches 114; Indels 60; Gaps 16;

QY 57 TEPEMLRNSTDTPL--TGFCIP--ESTTVEPAARRSTGLDAGGAVTELTLANMGNL 112
 Db 57 TTPMEEI--TTITTPMEEITITTPMVEETITLPMVAAMTTPMVEETITPTVETITTPMVEET 114
 QY 113 STDSAAAMEIQTTQAAA-----TEAQTPLAATEAQTTLTATEAQTPLAATEAQTTP 165
 Db 115 TITPMVEETITPMVEAMITLMEETMTPM-----EETITLPMVAAMTTPMVEET--TITTP 169

QY 166 PAATEAQTPTGLEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPA 225
 Db 170 MVEI--TTIPT-VETMT--PMVEAMTILP--MAAMTTP-----MBEETITTPMVEETIT 216
 QY 226 P-----EATEAQTPT-----ATEAQTPLAAMEALSTE-----PSATEALSMEP 266
 Db 217 PMVETITTPMVEAMT--PTEVVGSRMSVSTIRITTPMEAMITPTVETITLPTAAMTTPVEE 275
 QY 267 TTKRGLFIPFVSSTVTHKGIPTMAASNLVSVNYPVCGAPDHSVKQCLLAILLALV 320
 Db 276 TITPMVEETITPTVETITTPMVEAMITPTVETITTP-----MVLVMAVV 320

Search completed: July 14, 2003, 06:13:13
 Job time : 44 secs

A:Cross-references: SGD:S0003912
 A:Map position: 10R
 C:Keywords: transmembrane protein

Query Match 9.6%; Score 194; DB 2; Length 1161;
 Best Local Similarity 26.6%; Pred. No. 0.00012;
 Matches 73; Conservative 30; Mismatches 125; Indels 46; Gaps 11;

QY 56 ETEPEMLRNSTD---TTPLAG-----PCTP--ESTTVEPAARRSTGLDAGGAVT 100
 Db 62 EIVPSEIAAAVDFGDFTRLAGISGDEVTRMITGVPWYSTRKPAI--SSALSXKGIYV 119
 QY 101 EUTTELANMGNLSTDSAAAMEIQTTQPAATEAQTPLAATEAQTTLTATEAQTPLAATEA 160
 Db 120 AIPT---STSTTTKSTSTSTTTTITTTSTTTSTTTSTTTSTTTSTTTSTTTSTTTST 175
 QY 161 AQTTPAATEA---QTTQPTGLEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQT 215
 Db 176 TSTTPTTSTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 228

QY 216 TTAMEAQTTPAATEAQTTPAATEAQTPLA---MEALSTEP-----SATEALS 263
 Db 229 TPTTSTTTSTTSTTSTTSTTSTTSTTSTTPTTSTTPTTSTTPTTSTTPTTSTTPTTST 288
 QY 264 MPEPTTKRGLFIPFVSSTVTHKGIPTMAASNLVSNY 297
 Db 289 TAPTPTST--TSTTPTSTSSASASSVISTATTTITF 321

RESULT 14
 T21460
 hypothetical protein ZK945.10 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T21460
 R:Wilkinson, J.
 submitted to the EMBL Data Library, March 1995
 A:Reference number: Z19425
 A:Accession: T21460
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-796 <WIL>
 A:Cross-references: EMBL:Z48582; PIDN:CAA88469.1; GSPDB:GN00020; CESP:ZK945.10
 A:Experimental source: clone F27B5
 C:Genetics:
 A:Gene: CESP:ZK945.10
 A:Map position: 2
 A:Introns: 62/3; 124/3; 169/1; 216/1; 262/3; 371/1; 581/3; 608/1

Query Match 9.5%; Score 193.5; DB 2; Length 796;
 Best Local Similarity 22.6%; Pred. No. 8.6e-05;
 Matches 87; Conservative 57; Mismatches 170; Indels 71; Gaps 12;

QY 65 NSIDIT---PLTGGTPESTTVEPAARRSTGLDAGGAVTELTLANMGN-----LSTD 115
 Db 336 SSTSTTQQSSSITSSPSSITLSTIPTTT-----TPEITSTLSSLPDINAICSYLDET 388

QY 116 SAAMEIQTTQ---PAATEAQTPLAATEAQTTLTATEAQTPLAATEAQTTPAATEA 171
 Db 389 TITSTFTTITMLSTTTEESTTSTTTTTEVTSSTVITTEPTTLLTSTASTTTEPST 448

QY 172 QTTQPTGLEAQTTPAAMEAQT--TAPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPA 229
 Db 449 VITSPSTSPVSTVITSSSSSTVITPTTSTSTSPSS--TVITSTTAPSTSTGPPSS 506

QY 230 EAQTTPAATEAQTTPAAMEALSTEPSATEALSMEPTTKRGLFIPF--SVSSVTHKGIPTM 288
 Db 507 SETPSSASSVSTSSASSTQSSSTTQSSSTTQSSSTTQSSSTTQSSSTTQSSSTTQSSST 566
 QY 289 AASNLVSNYPVCGAPDHSVKQCLLAILLALVATIFFVCTVVLAVLRLSRKGMYP----- 343
 Db 567 TSVNLTNLSGLIGIYQTSIECTSTSSNVSVITKDGAC--FTKSVSMRPLGGTYPASTFV 625

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OR protein - protein search, using sw model

Run on: July 14, 2003, 06:10:58 ; Search time 23 Seconds
(without alignments)
724,934 Million cell updates/sec

Title: US-09-825-580-2

Perfect score: 2030

Sequence: 1 MFLQULLLLILLLGPGNSLQL.....TPPREDRERGGDILLHSFLRP 402

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt 40.4*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	2015	99.3	412	1 SEPL_HUMAN	Q14442 homo sapien
2	825.5	40.7	397	1 SEPL_MOUSE	Q62170 mus muscula
3	288.5	14.2	307	1 SGG3_DROME	P02840 drosophila
4	250.5	12.3	797	1 VGXK_HSVB	P28968 equine herp
5	196.5	9.7	5179	1 MUC7_HUMAN	Q02817 homo sapien
6	194	9.6	1161	1 DAN4_YEAST	P47179 saccharomyc
7	193.5	9.5	3178	1 YS89_CAEEL	Q09624 caenorhabdi
8	187	9.2	1367	1 AMYH_YEAST	P08640 saccharomyc
9	184	9.1	662	1 MUC1_XENLA	Q05049 xenopus lae
10	180.5	8.9	400	1 MUA1_XENLA	P10667 xenopus lae
11	179.5	8.8	328	1 SGG3_DROER	P13730 drosophila
12	176.5	8.7	699	1 VGJG_HSV2H	P13390 herpes simp
13	173.5	8.5	860	1 CH1Z_COCCIM	P54197 coccidioidie
14	171	8.4	5376	1 ZAN_MOUSE	O88799 mus musculu
15	169.5	8.3	1802	1 HKR1_YEAST	P41809 saccharomyc
16	168	8.3	1637	1 MRSP_STRAU	P80544 staphylococ
17	167.5	8.3	886	1 VGP3_BHAV6	Q07284 epstein-bar
18	167	8.2	405	1 SDG3_CHICK	P26261 gallus galli
19	167	8.2	605	1 YH8_YEAST	P38739 saccharomyc
20	166	8.2	528	1 POD4_HUMAN	O00592 homo sapien
21	164.5	8.1	2109	1 PGCA_CHICK	P07898 gallus galli
22	164	8.1	338	1 SEDI_YEAST	O01589 saccharomyc
23	163.5	8.1	1045	1 GUNB_CELFI	P26225 cellulomona
24	163	8.0	567	1 CH13_CANAL	P40954 candida alb
25	162.5	8.0	450	1 CY1_PARDE	P13627 paracoccus
26	162.5	8.0	670	1 VG50_HSV1A	Q00130 icetaurid h
27	161	7.9	676	1 MUC1_MESAU	O60528 mesocricetu
28	160.5	7.9	442	1 SDG3_RAT	P33671 rattus norv
29	160.5	7.9	1199	1 N121_RAT	P54511 rattus norv
30	159.5	7.9	442	1 SDG3_MOUSE	O64519 mus musculu
31	159.5	7.9	907	1 VGP3_BEV	F03200 epstein-bar
32	159	7.8	1758	1 Y1R7_YEAST	P40434 saccharomyc
33	159	7.8	1758	1 YJW5_YEAST	P40889 saccharomyc

RESULT 1	ID	SEPL_HUMAN	STANDARD	PRT	412 AA.	Q61548 mus musculu
AC	Q14242; Q12775;					P45386 haemophilus
DT	01-NOV-1997 (Rel. 35, Created)					Q28983 sus scrofa
DT	01-NOV-1997 (Rel. 35, Last sequence update)					O05140 rattus norv
DT	16-OCT-2001 (Rel. 40, Last annotation update)					O75056 homo sapien
DE	P-selectin glycoprotein ligand 1 precursor (PSGL-1) (Selectin P ligand) (CD162 antigen).					P08918 human cytom
GN	SELPLG.					P35658 homo sapien
OS	Homo sapiens (Human).					P13728 drosophila
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					P13595 mus musculu
OC	Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.					P03764 bacteriopho
OX	NCBI_TaxID=9606;					P17437 xenopus lae
RN	[1]					P39712 saccharomyc
RP	SEQUENCE FROM N.A.					
RC	TISSUE=Placenta;					
RX	MEDLINE=95332364; PubMed=7541799;					
RA	Veldman G.M., Bean K.M., Cumming D.A., Eddy R.L., Sait S.N.J., Shows T.B.;					
RT	"Genomic organization and chromosomal localization of the gene encoding human P-selectin glycoprotein ligand.";					
RL	J. Biol. Chem. 270:16470-16475(1995).					
RN	[2]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=9408493; PubMed=7505206;					
RA	Sako D., Chang X.J., Barone K.M., Vachino G., White H.M., Shaw G., Veldman G.M., Bean K.M., Cumming D.A., Furie B., Cumming D.A., Larsen G.R.,;					
RT	"Expression cloning of a functional glycoprotein ligand for P-selectin.";					
RL	Cell 75:1179-1186(1993).					
RN	[3]					
RP	SULFATION.					
RX	MEDLINE=96028106; PubMed=7585949;					
RA	Sako D., Comess K.M., Barone K.M., Camphausen R.T., Cumming D.A., Shaw G.D.;					
RT	"A sulfated peptide segment at the amino terminus of PSGL-1 is critical for P-selectin binding.";					
RL	Cell 83:323-331(1995).					
RN	[4]					
RP	SULFATION.					
RX	MEDLINE=96028107; PubMed=7585950;					
RA	Pouyani T., Seed B.;					
RT	"PSGL-1 recognition of P-selectin is controlled by a tyrosine sulfation consensus at the PSGL-1 amino terminus.";					
RL	Cell 83:333-343(1995).					
RN	[5]					
RP	SULFATION.					
RX	MEDLINE=96032682; PubMed=7559387;					
RA	Wilkins P.P., Moore K.L., McEever R.P., Cummins R.D.;					
RT	"Tyrosine sulfation of P-selectin glycoprotein ligand-1 is required for high affinity binding to P-selectin.";					
RL	J. Biol. Chem. 270:22677-22680(1995).					
CC	-1- FUNCTION: BINDS TO P-, E- AND L-SELECTINS. THE CALCIUM-DEPENDENT HIGH AFFINITY INTERACTION WITH P-SELECTIN MEDIATES THE TETHERING					

CC AND ROLLING OF NEUTROPHILS AND T-LYMPHOCYTES ON ENDOTHELIAL CELLS.
 CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
 CC -1- SUBCELLULAR LOCATION: TYPE I membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED ON NEUTROPHILS, MONOCYTES AND MOST
 CC LYMPHOCYTES. O-GLYCOSYLATED. ALSO N-GLYCOSYLATED.
 CC -1- PFM: HEAVILY O-GLYCOSYLATED.
 CC -1- PFM: SUBUNITED IN THE N-TERMINAL REGION; SULFATION IS NECESSARY
 CC FOR P-SELECTIN BINDING.
 CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD162 entry;
 CC WWW=http://www.ncbi.nlm.nih.gov/prow/cd/cd162.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U25956; AAA74577.1; -
 CC DR EMBL; U02297; AAC50061.1; -
 CC DR GlycoSubstEDB; O14242; -
 CC DR GeneW; HGNC;10722; SELPLG.
 CC DR MIM; 600738; -
 CC KW Cell adhesion; Glycoprotein; Transmembrane; signal; Repeat;
 CC Polymorphism; Sulfation.
 CC FT SIGNAL 1 17 POTENTIAL.
 CC FT PROPEP 48 41
 CC FT CHAIN 18 412 P-SELECTIN GLYCOPROTEIN LIGAND 1.
 CC FT DOMAIN 18 320 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 321 341 POTENTIAL.
 CC FT DOMAIN 342 412 CYTOPLASMIC (POTENTIAL).
 CC FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT DOMAIN 122 261 12 X 10 AA TANDEM REPEATS.
 CC FT REPEAT 122 131 1.
 CC FT REPEAT 132 141 2.
 CC FT REPEAT 142 151 3.
 CC FT REPEAT 162 171 4.
 CC FT REPEAT 182 191 5.
 CC FT REPEAT 192 201 6.
 CC FT REPEAT 202 211 7.
 CC FT REPEAT 212 221 8.
 CC FT REPEAT 222 231 9.
 CC FT REPEAT 232 241 10.
 CC FT REPEAT 242 251 11.
 CC FT REPEAT 252 261 12.
 CC FT MOD RES 46 46 SULFATION (POTENTIAL).
 CC FT MOD RES 48 48 SULFATION (POTENTIAL).
 CC FT MOD RES 51 51 SULFATION (POTENTIAL).
 CC FT VARIANT 132 141 MISSING (IN SHORT FORM; NOT AN
 CC ALTERNATIVE SPLICING).
 CC FT SEQUENCE 412 AA; 43201 MW; A92A2A902DC9963A CRCK4;
 CC SO

Query Match 99.3%; Score 2015; DB 1; Length 412;
 Best Local Similarity 97.6%; Pred. No. 9.4e-114; Indels 10; Gaps 1;
 Matches 402; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
 QY 1 MPIQLLLLLLLLPNGSLQMDTWADAEKALGPLLARDRRQATEVEYLDYDFLPETTEP 60
 Db 1 MPIQLLLLLLLLPNGSLQMDTWADAEKALGPLLARDRRQATEVEYLDYDFLPETTEP 60
 QY 61 EMIRNSTDTPPLTGGPGPESTTVEPAARRSTGLDAGAVTETLTELAMNGULSTDSAME 120
 Db 61 EMIRNSTDTPPLTGGPGPESTTVEPAARRSTGLDAGAVTETLTELAMNGULSTDSAME 120
 QY 121 IOTTOPAA-----TEAOTTPLAATAAOTTPLRLATAAOTTPLAATAAOTTPLAATE 170
 Db 121 IOTTOPAAATAAOTTOPVPTTAOTTPLAATAAOTTPLRLATAAOTTPLAATAAOTTPLAATE 180
 QY 121 AOTTOPGAEAOQTAPAAAEAOQTAPAAAEAOQTAPAAAEAOQTAPAAAEAOQTAPAAE 230

Db 181 AOTTOPGAEAOQTAPAAAEAOQTAPAAAEAOQTAPAAAEAOQTAPAAAEAOQTAPAAE 240
 QY 231 AOTTOPGAEAOQTAPAAAEAOQTAPAAAEAOQTAPAAAEAOQTAPAAAEAOQTAPAAE 290
 Db 241 AOTTOPGAEAOQTAPAAAEAOQTAPAAAEAOQTAPAAAEAOQTAPAAAEAOQTAPAAE 300
 QY 291 SMLSVNYPVGAADPHISVKOCLLAIIILAVATIFVFCYVLAVRLSRKGMYPVNYSP 350
 Db 301 SMLSVNYPVGAADPHISVKOCLLAIIILAVATIFVFCYVLAVRLSRKGMYPVNYSP 360

QY 351 EWCYISSLLIPDGGEGPSATANGSLSKAKSPGLTPEPREDREDDDTLHSFLP 402
 Db 361 EWCYISSLLIPDGGEGPSATANGSLSKAKSPGLTPEPREDREDDDTLHSFLP 412

RESULT 2
 SEPL_MOUSE STANDARD; PRT; 397 AA.
 AC 062170;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE P-selectin glycoprotein ligand 1 precursor (PSGL-1) (selectin P ligand).
 GN SELPLG OR SELPL OR SELPL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RX MEDLINE=96320265; PubMed=8639776;
 RA Yang J., Galipeau J., Kozak C., Furie B.C., Furie B.;
 RT 'mouse p-selectin glycoprotein ligand-1: molecular cloning, chromosome localization, and expression of a functional p-selectin receptor.';
 RL Blood 87:4176-4186(1996).
 CC -1- FUNCTION: BINDS TO P-, E- AND L-SELECTINS. THE CALCIUM-DEPENDENT HIGH AFFINITY INTERACTION WITH P-SELECTIN MEDIATES THE TETHERING AND ROLLING OF NEUTROPHILS AND T-LYMPHOCYTES ON ENDOTHELIAL CELLS.
 CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- PFM: HEAVILY O-GLYCOSYLATED. ALSO N-GLYCOSYLATED (BY SIMILARITY).
 CC -1- PFM: SUBUNITED IN THE N-TERMINAL REGION; SULFATION IS NECESSARY FOR P-SELECTIN BINDING (BY SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X91144; CAA62583.1; -
 CC DR MGD; MG1:106689; Selpl.
 CC KW Cell adhesion; Glycoprotein; Transmembrane; signal; Repeat; Sulfation.
 CC FT SIGNAL 1 17 POTENTIAL.
 CC FT PROPEP 48 41 BY SIMILARITY.
 CC FT CHAIN 18 397 P-SELECTIN GLYCOPROTEIN LIGAND 1.
 CC FT DOMAIN 18 307 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 308 328 POTENTIAL.
 CC FT DOMAIN 329 397 CYTOPLASMIC (POTENTIAL).
 CC FT MOD RES 54 54 SULFATION (POTENTIAL).
 CC FT MOD RES 56 56 SULFATION (POTENTIAL).
 CC FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT DOMAIN 126 225 10 X 10 AA TANDEM REPEATS.
 CC FT REPEAT 126 135 1.
 CC FT REPEAT 136 145 2.

FT	REPEAT	146	155	3.
FT	REPEAT	156	165	4.
FT	REPEAT	166	175	5.
FT	REPEAT	176	185	6.
FT	REPEAT	186	195	7.
FT	REPEAT	196	205	8.
FT	REPEAT	206	215	9.
FT	REPEAT	216	225	10.
SO	SEQUENCE	397 AA;	41781 MW;	512B645FFB21E4B2 CRC64;
	Query Match		40.7%;	Score 825.5; DB 1; Length 397;
	Best Local Similarity	44.7%;	Pred. No. 9.6e-43;	
	Matches 104;	Conservative 51;	Mismatches 120;	Indels 69; Gaps 8;
QY	1	MPLQLLLILLIPGNSLQIMPTWADAEKALGPIIARDRRQAT-EYEVLYDYPPEPTEP	59	
DB	1	MSPSFLVLLIIGPGNSLQIQDPWGHETKEAPVHLRRRQVVGDDDEDEDPYVNTDP	60	
QY	60	PEMLKNSITDTPP-----TGPPTPSTTVEPAAKRSGLDAGGAVTELTT	104	
DB	61	PELTKVNTVVAHPELPTVVWLEBDSSTAGSSEKATEKIAVTPDADGTGTA-----	115	
QY	105	ELANMGNLSTDSAAAMEIQTTPAATBAQTTPLAATBAQTTPLAATBAQTTP	164	
DB	116	---VGMLSIDSA-----TQMSLISVEVTVQPHSTVEVTS	145	
QY	165	PPAATEAQTTPPTLEAQTTPAAMAQAQTTPAAMAQAQTTPAAMAQAQTTP	224	
DB	146	QAPAPMAEFTSQAPAPMAEFTSQAPAPMAEFTSQAPAPMAEFTSQAPAPMAEFTS	205	
QY	225	APBAPBAQTTPAATBAQTTPP---AAMEALSTEPATDALSMPTTKRGLFIPPSVS--	279	
DB	206	KPAPTAETSKPAPTAETSKPAPTAETSKPAPTAETSKPAPTAETSKPAPTAETSK	265	
QY	280	----SVTH-----KGIMMAASNLISVNVYVGAPEHISVKQCCLALIIIAVAATF	328	
DB	266	FLGSPSVTHLPDGLKGLIVTPENSPAPLTPGSSDILIPKQCLLIIILIASLAVTFLVCT	325	
QY	329	VVLAVALSRKGGMYPVRYNSPTEVMVCISSILLDPGEGEPATANGISKAKSPGLPPEP	388	
DB	326	VVLAVALSRKHTMYPVRYNSPTEVMCISSILLDPGEGAVTANGGLPKYQD--LKTREPSG	383	
QY	389	DREGGDILTHSFLP 402		
DB	384	DRDGDILTHSFLP 397		

RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q.X., Chen L.X., RA Brandon R.C., Rogers J.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D., RA Man K.H., Doyle C., Barker E.G., Helt G., Nelson C.R., Miklos G.L.G., RA Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., RA Ballew K.W., Basu P.V., Bertan B.P., Bhandari D., Bolshakov S., RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotlier P., RA Burris K.C., Busam D.A., Butler H., Cadelu E., Center A.A., Chandra I., RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., RA Dodson K., Dong L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W., RA Foster C., Gabrielian A.E., Garg N.S., Galbraith W.M., Glasser K., RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., RA Hostin D., Houston K.A., Howland T.U., Wei M.-H., Ibegwam C., RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., RA Liu X., Martel B., McIntosh T.C., McLeod M.P., McPherson D., RA Merkulov G., Milphina N.V., Mobarry C., Morris J., Moshrefi A., RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclib J.M., RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G., RA Palzer K., Remington K., Saunders R.D.C., Scheeler F., Shen H., RA Shue B.C., Siden-Klimos I., Simpson M., Skupski M.P., Smith T., RA Spiers E., Spradling A.C., Stapleton M., Strong R., Sun E., RA Svitak R., Tector C., Turner R., Venter E., Wang A.H., Wang X., RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J., RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; RA RT "The genome sequence of Drosophila melanogaster."; RA RL Science 287:2185-2195 (2000).
RP	SEQUENCE OF 1-28 FROM N.A.
RX	MEDLINE=88332966; PubMed=3138416;
RA	Martin C.H., Mayeda C.A., Meyerowitz E.M.;
RT	"Evolution and expression of the Sgs-3 glue gene of Drosophila.";
RL	J. Mol. Biol. 201:273-287 (1988).
RN	[4]
RP	DEVELOPMENTAL STAGE.
RX	MEDLINE=94038699; PubMed=8223281;
RA	Huet F., Ruiz C., Richards G.;
RT	"Puffs and PCR: the in vivo dynamics of early gene expression during ecdysose responses in Drosophila.";
RL	Development 118:613-627 (1993).
CC	-1- SUBCELLULAR LOCATION: Secreted.
CC	-1- TISSUE SPECIFICITY: SALIVARY GLAND SPECIFIC.
CC	-1- DEVELOPMENTAL STAGE: IN THE SALIVARY GLANDS OF MID INSTAR LARVAE LEVELS DRAMATICALLY INCREASE DURING PUFF STAGE 1 AT 98-106 HOURS OF DEVELOPMENT. LEVELS REMAIN CONSTANT AND ABUNDANT IN LATE LARVAE UNTIL PUFF STAGE 10, THEN DECREASE BY STAGE 11.
CC	This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sdb.ch/announce/ or send an email to license@sdb-ebi.ch).
CC	-----
DR	EMBL; X01918; CAA25594.1; .
DR	EMBL; AE003544; AAF50056.1; .
DR	EMBL; X78392; CAA55154.1; .
DR	PIR; A03329; GSPF3.
DR	FLYBase; FBgn003373; Sgs3.
KW	Repeat; Signal.
KM	POTENTIAL.
FT	SIGNAL 1 23
FT	CHAIN 24 307 SALIVARY GLUE PROTEIN Sgs-3.

SQ SEQUENCE 307 AA; 32196 MW; 45803DEDI6C418BC CRC64;
 Query Match 14.2%; Score 288.5; DB 1; Length 307;
 Best Local Similarity 30.8%; Pred. No. 8,4e-11;
 Matches 69; Conservative 32; Mismatches 112; Indels 11; Gaps 2;
 QY 55 PEPPEPMLNNSDTDTPLTGGPESSTVEPA-----RSTGADAGAVTELTEL 106
 DB 37 PRTLPQPCPTTTTTCAPRQOSTQOPPCPTTSKPTTPKQTLTQLP---CTTPPTTKA 93
 QY 107 ANGNLSTDSAAAEIOTTGPAAPAEAOPTPLAAAEAOPTPLAAAEAOPTPLAAAEAOPTPL 166
 DB 94 TTTKPTTKATTTTAAVTKPTTKQTTLQPLCTTPPTTKQTLQPLCTTPPTTKPTTKP 153
 QY 167 AATEAQTOTPTGAEAOPTTAPPAEAOPTTAPPAEAOPTTAPPAEAOPTTAPPAEAOPTTAP 226
 DB 154 TTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKP 213
 QY 227 EAPEAOPTTPTPAEAOPTTAPPAEALSTEPSAEALSMETPKR 270
 DB 214 TTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTK 257

RESULT 4
 VGLX_HSVEB STANDARD; PRT; 797 AA.
 AC P28968;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-DEC-1992 (Rel. 24, Last annotation update)
 DE Glycoprotein X precursor.
 GN 71.
 OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirinae.
 OX NCBI_TaxID=31520;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92295566; Pubmed=1318606;
 RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
 RT "The DNA sequence of equine herpesvirus-1.";
 RL Virology 189:304-316(1992).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 CC EMBL; M86664; AAB02506.1; -
 DR PIR; H36802; VEBEX1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 797
 FT DOMAIN 23 465
 FT TRANSMEM 766 790
 FT CARBOHYD 590 590
 FT SEQUENCE 797 AA; 80342 MW; 50C9BD9211F5B5E2 CRC64;

Query Match 12.3%; Score 250.5; DB 1; Length 797;
 Best Local Similarity 28.0%; Pred. No. 4.3e-08;
 Matches 104; Conservative 33; Mismatches 169; Indels 65; Gaps 13;

QY 55 PEPPEPMLNNSDTDTPLTGGPESSTVEPAARRSGLDAGAVTELTELPLANNGL 112
 DB 53 PRTSP-----TSSSPSTHTSSSPSTHTSSSPSTHTSSSPSTHTSSSPSTHTSSSPSTHTSS 106
 QY 113 STDSAAAEIOTTGPAAPAEAOPTTAPPAEAOPTTAPPAEAOPTTAPPAEAOPTTAP 155
 DB 107 ETTTTPPTASTTTPTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 166

QY 156 -----LAATEAQTOTPTGAEAOPTTAPPAEAOPTTAPPAEAOPTTAPPAEAOPTTAP 205
 DB 167 TTTAATTTVPTASTTTTDTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 226
 QY 206 PAAMEAQTOTPTGAEAOPTTAPPAEAOPTTAPPAEAOPTTAPPAEAOPTTAPPAEAOPTTAP 264
 DB 227 SAATAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 286
 QY 265 EPTTKRGLPFPVSSVTHKGIPTMA---SNLSVNVVPGAP--DHISVQCLLAILLIL 319
 DB 287 TPTS-----TSSAALTSTPTPTSAAATSAESTTEAAPTSTPTDPTTPE-----AT 333
 QY 320 VAITFVCTVLAVALRSRKHMYEVRNVPSTEMVCISSLDPGGSPATANGLSKAKS 379
 DB 334 TATTSPESTVSASTSA-----TTTAFTTESHSTSPDSSSTGASTAEPSSSTPTLT 383
 QY 380 PGLTPEPRDR 390
 DB 384 PS-TATPSTDQ 393

RESULT 5
 MUC2_HUMAN STANDARD; PRT; 5179 AA.
 AC 002817; Q14878;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Mucin 2 precursor (Intestinal mucin 2).
 GN MUC2 OR SMUC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Intestine;
 RX MEDLINE=94132002; Pubmed=8300571;
 RA Gum J.R., Hicks J.W., Toribara N.W., Siddiki B., Kim Y.S.;
 RT "Molecular cloning of human intestinal mucin (MUC2) cDNA.
 RT Identification of the amino terminus and overall sequence similarity
 RT to prepro-von Willebrand factor.";
 RL J. Biol. Chem. 269:2440-2446(1994).
 RN [2]
 RP SEQUENCE OF 626-1895 AND 4196-5179 FROM N.A.
 RC TISSUE=Colon;
 RX MEDLINE=93016075; Pubmed=1400449;
 RA Gum J.R., Hicks J.W., Toribara N.W., Roche E.-M., Lagace R.E.,
 RA Kim Y.S.;
 RT "The human MUC2 intestinal mucin has cysteine-rich subdomains located
 RT both upstream and downstream of its central repetitive region.";
 RL J. Biol. Chem. 267:21375-21383(1992).
 RN [3]
 RP SEQUENCE OF 1343-1895 AND 4176-4195 FROM N.A.
 RX MEDLINE=9158717; Pubmed=1885763;
 RA Toribara N.W., Gum J.R., Culhane P.J., Lagace R.E., Hicks J.W.,
 RA Petersen G.M., Kim Y.S.;
 RT "MUC-2 human small intestinal mucin gene structure. Repeated arrays
 RT and polymorphism.";
 RL J. Clin. Invest. 88:1005-1013(1991).
 CC -1- FUNCTION: COATS THE EPITHELIA OF THE INTESTINES, AIRWAYS, AND
 CC OTHER MUCOS MEMBRANE-CONTAINING ORGANS. THOUGHT TO PROVIDE A
 CC PROTECTIVE, LUBRICATING BARRIER AGAINST PARTICLES AND INFECTIOUS
 CC AGENTS AT MUCOSAL SURFACES.
 CC -1- SUBUNIT: MULTIMERIC.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: COLON, SMALL INTESTINE, COLONIC TUMORS,
 CC BRONCHUS, CERVIX AND GALL BLADDER.
 CC -1- PTM: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR
 CC INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).
 CC -1- POLYMORPHISM: THE NUMBER OF REPEATS IS HIGHLY POLYMORPHIC AND
 CC VARIES AMONG DIFFERENT ALLELES.
 CC -1- SIMILARITY: THE N-TERMINAL DOMAIN SHOWS SOME SIMILARITY TO THAT

OF SILKORM HEMOCYTIN, CONTAINS 2 WFPC DOMAINS.
 -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
 -2- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: L21998; AAB95295.1; -
 EMBL: M74027; AAA59875.1; -
 EMBL: M94131; AAA59163.1; -
 EMBL: M94132; AAA59164.1; -
 Genew: HGNC:7512; MUC2.
 MIM: 158370; -
 InterPro: IPR000359; Cys_knot.
 InterPro: IPR000561; EGF-like.
 InterPro: IPR002400; GF_cysknot.
 InterPro: IPR002919; TIL_Cystrich.
 InterPro: IPR001846; WFPC.
 InterPro: IPR00093; Cys_knot; 1.
 Pfam: PF00094; vwd; 4.
 Pfam: PF01826; TIL; 1.
 PRINTS: PR00438; GFCSYKNOT.
 SMART: SM00214; VMC; 2.
 SMART: SM00216; VMD; 4.
 PROSITE: PS00022; EGF_1; UNKNOWN_1.
 PROSITE: PS01185; CTCK_1; 1.
 PROSITE: PS01225; CTCK_2; 1.
 PROSITE: PS01208; WFPC_2.1.
 GlycoProfile: Repeat; Signal.
 GlycoProfile: Repeat; Signal.
 CHAIN 1 20
 DOMAIN 1401 1747
 REPEAT 1401 1416
 REPEAT 1417 1432
 REPEAT 1433 1448
 REPEAT 1449 1464
 REPEAT 1465 1471
 REPEAT 1472 1478
 REPEAT 1479 1494
 REPEAT 1495 1517
 REPEAT 1518 1533
 REPEAT 1534 1556
 REPEAT 1557 1572
 REPEAT 1573 1596
 REPEAT 1597 1612
 REPEAT 1613 1635
 REPEAT 1636 1651
 REPEAT 1652 1675
 REPEAT 1676 1683
 REPEAT 1684 1699
 REPEAT 1700 1715
 REPEAT 1716 1731
 REPEAT 1732 1747
 DOMAIN 4815 4886
 DOMAIN 4924 4991
 DOMAIN 5075 5160
 DISULFID 5089 5136
 DISULFID 5098 5152
 DISULFID 5154 5154
 DISULFID 5159 ?
 CARBOHYD 163 163
 CARBOHYD 423 423
 CARBOHYD 670 670
 CARBOHYD 770 770
 CARBOHYD 894 894

FT CARBOHYD 1139 1139 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1154 1154 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1215 1215 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1230 1230 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1246 1246 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1267 1267 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1820 1820 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4339 4339 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4351 4351 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4362 4362 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4373 4373 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4422 4422 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4438 4438 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4502 4502 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4616 4616 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4627 4627 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4752 4752 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4787 4787 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4881 4881 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4888 4888 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4955 4955 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4970 4970 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 5019 5019 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 5038 5038 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 5069 5069 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 1351 1351 H -> L (IN REF. 3).
 FT CONFLICT 1412 1412 T -> S (IN REF. 3).
 FT CONFLICT 1449 1449 L -> P (IN REF. 3).
 FT CONFLICT 1504 1504 M -> T (IN REF. 3).
 FT CONFLICT 4192 4192 G -> S (IN REF. 2).
 SQ SEQUENCE 5179 AA; 540295 MW; 85CD7571FB9A5663 CRC64;

Query Match
 Best Local Similarity 9.7%; Score 196.5; DB 1; Length 5179;
 Matches 89; Conservative 24; Mismatches 169; Indels 81; Gaps 13;

OY 55 PTEPPEMLRNSTDTPTLTPGPTPESTVVEPARRSTGIDAGAVTELTTELANNGNIST 114
 Db 3966 PTTTITTTTTVTPPTPTGQTPTTTTITTTTTVTPPTPTGQTPTTTTITTTTTV- 4024

OY 115 DSAMEIQTOPAARE-QTTPLAATEAQTTLTTEAO--TTELAATEAQTTPAATE 170
 Db 4025 -----PPTPTPTGQTPTTTTITTTTTVTPPTPTGQTPTTTTITTTTTVTPPTPTG 4077

OY 171 AQTPTGAEAQT-----APAAAEAQTAPAAEAQTTP-PAEAEAQTOTTAAEAQT 224
 Db 4078 TQTPPTPTPTTTTTVTPPTPTPTGQTPTTTTITTTTTVTPPTPTGQTPTTTTITTTTT 4137

OY 225 -----APEATEAQTTPATEAQTTPLAAMEALSTEPSATEALSMETTKRGLTFPFSV 279
 Db 4138 VTPPTPTPTGQTPTTTTITTTTTVTPPTPTPTGQTPTTTTITTTTTVTPPTPTPTPTPT 4190

OY 280 SVTHKGIEMAS-----NISVNYPVGAPDHSVKQCLAILLALVATTFVVCVAVLAVR 334
 Db 4191 TGTGTGPPPTHTSTABIABLITSNP--PPEASTPO----- 4222

OY 335 LSRKGMVPRVRYSP-TEWVCISLPPDGGKSPS-----ATANGGSKAKSPGTT 383
 Db 4223 TSKS-----TSSPLTESITLSTLPPAIEWTSTAPPSTPTTPTTSGGHTLSPSPSTT 4275

OY 384 PEP 386
 Db 4276 TSP 4278

RESULT 6
 DN4 YEAST STANDARD; PRT; 1161 AA.
 AC P47179;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cell wall protein DN4 precursor.

GN DAN4 OR YJRI51C OR J2223.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RA SEQUENCE FROM N.A.
 RP Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP REGULATION.
 RX MEDLINE=21113166; Pubmed=11160904;
 RA Cohen B.D., Serfil O., Abramova N.E., Davies K.J., Lowry C.V.;
 RT "Induction and repression of DAN1 and the family of anaerobic
 RT manoprotein genes in Saccharomyces cerevisiae occurs through a
 RT complex array of regulatory sites";
 RL Nucleic Acids Res. 29:799-808(2001).
 CC -1- FUNCTION: COMPONENT OF THE CELL WALL (By similarity).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
 CC (Potential).
 CC -1- PTM: EXTENSIVELY O-GLYCOSYLATED (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE SRP1 / T1P1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z49651; CAAB9684.1; -
 DR SGI; S0003912; DAN4.
 DR InterPro; IPR000992; SRP1_T1P1.
 DR Pfam; PF00660; SRP1_T1P1.1.
 DR PROSITE; PS00724; SRP1_T1P1, 1.
 DR Cell wall; Glycoprotein; Membrane; GPI-anchor; signal.
 KW CHAIN 1 24 POTENTIAL.
 FT CHAIN 25 1146 CELL WALL PROTEIN DAN4.
 FT PROPEP 1147 1161 REMOVED IN MATURE FORM (POTENTIAL).
 FT LIPID 1146 1146 GPI-ANCHOR (POTENTIAL).
 FT SEQIDCE 1161 AA; 118358 MW; 7954CT5D69F0CA58 CRC64;
 SQ
 Query Match 9.6%; Score 194; DB 1; Length 1161;
 Best local Similarity 26.6%; Pred. No. 0.00015;
 Matches 73; Conservative 30; Mismatches 125; Indels 46; Gaps 11;

DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DR Hypothetical protein ZK945.9 in chromosome II.
 GN ZK945.9/ZK945.10
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Peloderae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RA SEQUENCE FROM N.A.
 RP STRAIN=Briceol N2;
 RC Wilkinson-Sprat J.;
 RA Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP REVISIONS.
 RA Durbin R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: CONTAINS 1 GPS DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z48544; CAB70192.1; -
 DR EMBL; Z48582; CAB70192.1; JOINED.
 DR EMBL; Z48544; CAB70201.1; -
 DR WormRep; ZK945.9; CE25697.
 DR InterPro; IPR002111; Cat channel_TripL.
 DR InterPro; IPR001024; Lipoxigenase_LH2.
 DR InterPro; IPR000636; M+channel_nlg.
 DR Pfam; PF00520; Ion trans; 1.
 DR Pfam; PF01477; PLAT; 1.
 DR Pfam; PF01825; GPS; 1.
 DR SMART; SM00303; GPS; 1.
 DR SMART; SM00308; LH2; 1.
 DR Hypothetical protein; Transmembrane.
 KW Hypothetical protein; Transmembrane.
 FT DOMAIN 266 1196 SER/THR-RICH.
 FT DOMAIN 1105 1241 GLY/SER-RICH.
 FT DOMAIN 2071 2120 GPS.
 FT TRANSMEM 13 30 POTENTIAL.
 FT TRANSMEM 51 73 POTENTIAL.
 FT TRANSMEM 2139 2161 POTENTIAL.
 FT TRANSMEM 2348 2367 POTENTIAL.
 FT TRANSMEM 2390 2412 POTENTIAL.
 FT TRANSMEM 2451 2468 POTENTIAL.
 FT TRANSMEM 2483 2505 POTENTIAL.
 FT TRANSMEM 2567 2589 POTENTIAL.
 FT TRANSMEM 2836 2858 POTENTIAL.
 FT TRANSMEM 2939 2961 POTENTIAL.
 FT TRANSMEM 2976 2998 POTENTIAL.
 FT TRANSMEM 3038 3060 POTENTIAL.
 SQ SEQIDCE 3178 AA; 344726 MW; F8239436D03666CD CRC64;
 Query Match 9.5%; Score 193.5; DB 1; Length 3178;
 Best local Similarity 22.6%; Pred. No. 0.00048;
 Matches 87; Conservative 57; Mismatches 170; Indels 71; Gaps 12;

RESULT 7
 YS89 CABEL
 ID YS89 CABEL STANDARD; PRT; 3178 AA.
 AC Q09624; Q09625; Q969D4;
 DT 01-NOV-1995 (Rel. 32, Created)

Db 449 VTTSPSTPVTSTVTSSTSSSTVTTVTPTSTSTSPSS--TVTSTTSPSTSTSPSS 506
 Qy 230 BAQTOPTATBAQTPPLAMEALSTEPSATMALSMPTTKRGLFIFP-SVSSVTHKIGIM 288
 Db 507 SSTPSTASSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 566
 Qy 289 MASNLSVNYVPGAPPHISVKQCLLALLLVALVATTFVCTVVALVAVLSRKHGMYR----- 343
 Db 567 TSVNLTLLNGGLGIGVQTSIECTSPSSNMYVSTTKQAC-FTKSVEMRPLDGGTYYPASTV 625
 Qy 344 -----VKNYSPTEWCISSLLPDDGEGPSAIPANGLSRA 377
 Db 626 GPGNYTPRATWTTDDKRVYTYVAVNYIOEYSSSTTISESS-----TSVAASSSTSTP 677
 Qy 378 KSPGLT-----PEPEDEREGDGLT 396
 Db 678 STPSSSTLSTSTVTEBSTRSSSDSTT 702

RESULT 8

AMWH YEAST STANDARD; PRT; 1367 AA.
 ID AMWH YEAST STANDARD; PRT; 1367 AA.
 AC P08640; P08068; Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 31, Last sequence update)
 DT 01-FEB-1995 (Rel. 40, Last annotation update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE GLUCAMYLASE S1/S2 precursor (BC 3.2.1.3) (Glucan 1,4-alpha-glucosidase) (1,4-alpha-D-glucan glucanhydrolase).
 GN ST1A OR ST1A2 OR MAL5 OR Y1R019C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OC NCB1_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288c / AB972;
 RA Barrett B.G., Badcock K., Bankier A.T., Bowman S., Brown D., Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A., Gentles S., Hamlyn N., Hornsell T.S., Hunt S., Jagels K., Jones M., Louis E., Lye G., Mould S., Mould T., Odell C., Pearson D., Rajadream M.A., Riles L., Rowley N., Skelton J., Smith V., Walsh S.V., Whitehead S.;
 RA Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RN SEQUENCE OF 1-242 AND 762-1331 FROM N.A.
 RP MEDLINE=87194600; PubMed=3106330;
 RA Yamashita I., Nakamura M., Fukui S.;
 RT "Gene fusion is a possible mechanism underlying the evolution of ST1A.";
 RL J. Bacteriol. 169:2142-2149(1987).
 RN [3]
 RP SEQUENCE OF 1-31 FROM N.A.
 RC STRAIN=SPX101-1C;
 RA MEDLINE=89031330; PubMed=3141213;
 RA Pardo J.M., Ianez E., Zalacain M., Claros M.G., Jimenez A.;
 RT "Similar short elements in the 5' regions of the STR2 and SGA genes from Saccharomyces cerevisiae.";
 RL FBSB Lett. 239:1175-1184(1988).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-glucose residues successively from non-reducing ends of the chains with release of beta-D-glucose.
 CC -1- SIMILARITY: TO S.POMBE SPC215.13.
 CC -1- SIMILARITY: SOME, TO S.POMBE SPC285.13C.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; 238061; CA86176.1; -
 DR EMBL; M16164; AAA35014.1; -
 DR EMBL; M16165; AAA35015.1; -
 DR EMBL; X13857; CAA32069.1; -
 DR PIR; B26877; B26877.
 DR PIR; A26877; A26877.
 DR PIR; S48478; S48478.
 DR SGD; S0001458; M0C1.
 KW Hydrolyase; Glycosidase; Polysaccharide degradation; Glycoprotein;
 KW Signal; Multigene family.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 1367 GLUCAMYLASE S1/S2.
 FT DOMAIN 210 1367 SER/THR-RICH.
 FT CARBOHYD 817 817 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 874 874 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1367 AA; 136110 MW; 91C00E2B9D51AA9D CRC64;

Query Match

Best Local Similarity 24.0%; Pred. No. 0.00048;
 Matches 55; Conservative 36; Mismatches 110; Indels 28; Gaps 4;
 Qy 76 GPESTVTPRARRSTGIDAGAVTELTTELAMGNLSIPDSAMEIQTTQAPATQAQTPR 135
 Db 207 GKSSSTTSSSTSSST-----TSSSTSSSTTSSSTSSSTT 244
 Qy 136 LAATBAQTRRLTATBAQTPPLAATBAQTPPAPATBAQTPPGLBAQTPAPAMBAQTPA 195
 Db 245 SSTSSSTSSSTTATATPTTSCVKEKPTPTTSCVKEKPTPTTSCVKEKPTPTTSCVKEKPTTTS 303
 Qy 196 PAMBAQTPPPAPAMBAQTPPTTAPAMBAQTPAPATBAQTPPTA--TEAQTPLAAMEALS 253
 Db 304 KTCCKTKTTPVPTPSSSTSSSAPVPTPSSSTSSSAPVPTPSSSTSSSAPVPTPSSST 363
 Qy 254 TEPASATMALSMPTTKRGLFIFPSSVTHKGIIPMAASLSTNYVVGAP 302
 Db 364 TESSSAPVTS--STSSSAPVTSSTSSSAPVPTPSSSTSSSAP 409

RESULT 9

M0C1_XEN1A STRAND; PRT; 662 AA.
 ID M0C1_XEN1A STRAND; PRT; 662 AA.
 AC Q05049;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Integumentary mucin C.1 (FIM-C.1) (Fragment).
 OS Xenopus laevis (African clawed frog).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Duteleostomi; OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCB1_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4; 5; 6 AND 7).
 RC TISSUE=SKIN;
 RA MEDLINE=93077556; PubMed=1447205;
 RA Hauser F., Hoffmann W.;
 RT "p-domains as shuffled cysteine-rich modules in integumentary mucin C.1 (FIM-C.1) from Xenopus laevis. Polydispersity and genetic polymorphism.";
 RL J. Biol. Chem. 267:24620-24624(1992).
 CC -1- FUNCTION: COULD BE INVOLVED IN DEFENSE AGAINST MICROBIAL INFECTIONS. PROTECTS THE EPITHELIA FROM EXTERNAL ENVIRONMENT.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- ALTERNATIVE PRODUCTS: At least 7 isoforms; 1 (shown here), 2, 3, 4, 5, 6 and 7; may be produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: SKIN.
 CC -1- PTM: EXTENSIVELY O-GLYCOSYLATED.
 CC -1- SIMILARITY: CONAINS 6 P-TYPE (TRPFLD) DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sb-sib.ch).

EMBL: L02115: AAA74725.1; -

DR PIR; A45155; A45155.

DR HSP; P01359; 2PSP.

DR InterPro: IPR000519; P_trefol1.

DR SMART; SMO0018; trefol1; 6.

DR PROSITE: PS00025; P_TREFOL1; 6.

DR Repeat; Amphibian skin; Glycoprotein; Alternative splicing.

KW NON_TER

FT DOMAIN 1 144

FT REPEAT 81 88

FT REPEAT 89 96

FT REPEAT 97 104

FT REPEAT 105 112

FT REPEAT 113 120

FT REPEAT 121 128

FT REPEAT 129 136

FT REPEAT 137 144

FT DOMAIN 161 202

FT DOMAIN 218 301

FT REPEAT 218 224

FT REPEAT 225 239

FT REPEAT 240 249

FT REPEAT 250 259

FT REPEAT 260 275

FT REPEAT 276 287

FT REPEAT 288 294

FT REPEAT 295 301

FT DOMAIN 306 347

FT DOMAIN 353 394

FT DOMAIN 402 522

FT REPEAT 402 411

FT REPEAT 412 419

FT REPEAT 420 431

FT REPEAT 432 443

FT REPEAT 444 453

FT REPEAT 454 460

FT REPEAT 461 472

FT REPEAT 473 480

FT REPEAT 480 491

FT REPEAT 492 498

FT REPEAT 499 515

FT REPEAT 516 522

FT DOMAIN 525 566

FT DOMAIN 572 613

FT DOMAIN 620 661

FT DISULFID 162 188

FT DISULFID 172 187

FT DISULFID 182 199

FT DISULFID 307 333

FT DISULFID 317 332

FT DISULFID 327 344

FT DISULFID 354 380

FT DISULFID 364 379

FT DISULFID 374 391

FT DISULFID 526 551

FT DISULFID 536 552

FT DISULFID 546 563

FT DISULFID 573 599

FT DISULFID 583 598

FT DISULFID 593 610

FT DISULFID 621 647

FT DISULFID 631 646

FT DISULFID 641 658

FT VARSPLIC 240 259

FT VARSPLIC 256 294

8 X 8 AA APPROXIMATE TANDEM REPEATS, ALN/THR-RICH.

1-1.

1-2.

1-3.

1-4.

1-5.

1-6.

1-7.

1-8.

P-TYPE 1.

8 X APPROXIMATE TANDEM REPEATS, THR-RICH.

2-1.

2-2.

2-3.

2-4.

2-5.

2-6.

2-7.

2-8.

P-TYPE 2.

P-TYPE 3.

12 X APPROXIMATE TANDEM REPEATS, THR-RICH.

3-1.

3-2.

3-3.

3-4.

3-5.

3-6.

3-7.

3-8.

3-9.

3-10.

3-11.

3-12.

P-TYPE 4.

P-TYPE 5.

P-TYPE 6.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

MISSING (IN ISOFORMS 5 AND 6).

MISSING (IN ISOFORMS 4 AND 7).

MISSING (IN ISOFORMS 4 AND 6).

FT VARSPLIC 278 278 MISSING (IN ISOFORM 5).

FT VARSPLIC 306 350 MISSING (IN ISOFORMS 2 AND 5).

FT VARSPLIC 420 498 MISSING (IN ISOFORM 3).

FT VARIANT 276 276 K -> E.

FT VARIANT 354 354 C -> R.

FT VARIANT 415 415 T -> A.

SQ SEQUENCE 662 AA; 67774 MW; F085277F1ED2FD40 CRC64;

Query Match 9.1%; Score 184; DB 1; Length 662;

Best Local Similarity 27.5%; Pred. No. 0.00033;

Matches 80; Conservative 20; Mismatches 111; Indels 80; Gaps 13;

QY 57 TEPPEMLRNSDTPPTLTGSPPESTT---VEPARRSTGLDAGAVTE----- 101

DB 278 TTTPTTTTTTTTPTTTTKATTTTTTSGCKMPEPSRREDQGY---SGIIESOCRTKGCDF 334

QY 102 -----LTTTELAMNGNLTSDSAAHEIQ----- 122

DB 335 DSSIPQTKWCFYTLGQVADCKVEPSQRVDCGFRGITADDCRQKNCDFSSISGTRKWCFFYS 394

QY 123 TTPPAAETEAQTTPLAATEAQTTRLTATEAQTTPPLAATEAQTTPPAATEAQTTPGLEAQ 182

DB 395 TQVAAATKTTTPTTTTPTTTT---TTTKATTTTTPTTT---TTTTTTTKAF 446

QY 183 TTPAAPAAEAQTTPAAPAAEAQTTPPAAMEAQTTPQTTAMEAQTTPAPEATEAQTTPATEAQ 242

DB 447 TTTPTTTTTPTTTTKA---TTTTPTTTTTPTTTT---KATTTTTPTTT---TTTTPTTTTK 499

QY 243 ---TTPPLAAMEALSTEPSATTEALS---MEPTTKGLFIP-RSVSSVTHKG 285

DB 500 ATTTPTTTTTTTTTTTTKATTTTTTSGCKMPEPSRKRADCGYGITESQCRSKG 550

RESULT 10

MDAL_XENLA STANDARD; PRT; 400 AA.

ID MDAL_XENLA

AC P10667;

DT 01-JUL-1989 (Rel. 11, Created)

DT 01-JUL-1989 (Rel. 11, Last sequence update)

DE 01-OCT-1994 (Rel. 30, Last annotation update)

DE Integumentary mucin A.1 precursor (FIM-A.1) (Preprospasmodysin).

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus.

OX NCBI_TaxID=8355;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=88227968; PubMed=3372504;

RA Hoffmann W.;

RT "A new repetitive protein from Xenopus laevis skin highly homologous to pancreatic spasmodic polypeptide.";

RL J. Biol. Chem. 263:7686-7690(1988).

RN [2]

RP CHARACTERIZATION.

RX MEDLINE=90316191; PubMed=2196180;

RA Hauser F., Gerzsen E.M., Hoffmann W.;

RT "Expression of spasmodysin (FIM-A.1): an integumentary mucin from Xenopus laevis.";

RL Exp. Cell Res. 189:157-162(1990).

CC -1- FUNCTION: COULD BE INVOLVED IN DEFENSE AGAINST MICROBIAL INFECTIONS. PROTECTS THE EPITHELIA FROM EXTERNAL ENVIRONMENT.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: EXPRESSED AND STORED EXCLUSIVELY IN MATURE MUCOUS GLANDS OF THE SKIN.

CC -1- PTM: EXTENSIVELY O-GLYCOSYLATED. CONSIST OF ABOUT 70% CARBOHYDRATE AND 30% PROTEIN.

CC -1- SIMILARITY: CONTAINS 4 P-TYPE (TREFOL1) DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL; M19971; AAA49960.1; -
 DR PIR; A28172; A28172.
 DR HSSP; P01359; 2PSP.
 DR InterPro; IPR000519; P_crefc011.
 DR Pfam; PF000088; trefc011_4.
 DR PRINTS; PR00680; PTRREF01.
 DR SMART; SM00018; P_4.
 DR PROSITE; PS00025; P_TREF01L; 3.
 KW Repeat; Signal; Amphibian skin; Glycoprotein.
 FT SIGNAL 1 20
 FT CHAIN 1 20 POTENTIAL.
 FT DOMAIN 21 400 INTEGRMENTARY MUCIN A.1.
 FT DOMAIN 22 63 P-TYPE 1.
 FT DOMAIN 73 114 P-TYPE 2.
 FT DOMAIN 127 252 14 X 9 AA TANDEM REPEATS.
 FT DOMAIN 272 259 7 X 4 AA REPEATS OF E-T-T-T.
 FT DOMAIN 304 342 P-TYPE 3.
 FT DOMAIN 352 393 P-TYPE 4.
 FT DISULFID 23 49 BY SIMILARITY.
 FT DISULFID 33 48 BY SIMILARITY.
 FT DISULFID 43 60 BY SIMILARITY.
 FT DISULFID 74 100 BY SIMILARITY.
 FT DISULFID 84 99 BY SIMILARITY.
 FT DISULFID 94 111 BY SIMILARITY.
 FT DISULFID 312 337 BY SIMILARITY.
 FT DISULFID 322 339 BY SIMILARITY.
 FT DISULFID 353 379 BY SIMILARITY.
 FT DISULFID 363 378 BY SIMILARITY.
 FT DISULFID 373 390 BY SIMILARITY.
 FT CARBOHYD 63 63 N-LINKED (GLCNAC..)(POTENTIAL).
 SQ SEQUENCE 400 AA; 42641 MW; AC089F042713BCAB CRC64;

Query Match 8.9%; Score 180.5; DB 1; Length 400;
 Best Local Similarity 31.0%; Pred. No. 0.00031;
 Matches 67; Conservative 8; Mismatches 88; Indels 53; Gaps 9;

QY 57 TEPPEMLRNSDTPPLTGPSTPSTVPEPARRSTGLDAGAVTELTTELANMGNISDTS 116
 DB 119 TPAPDTTASTAETAETTTVPTTPTPTTV----- 145

QY 117 AAMEIQTPAATEAQTTPLAATEAQTTRITLA--TEAQTTPLAATEAQTTPPAATEAQT 174
 DB 146 -----PTTPETTTVPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 198

QY 175 OPTGLEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTP 234
 DB 199 VPTTPEET-ITVPTTPEET-TTVPTTPEET-TTVPTTPEET---AETTTVPTTPEETTT 252

QY 235 OPTAT-EAQTTPAAMEALSTEPSATREALSMERTTK 269
 DB 253 EPPTTPTTDTTP---PTLPPTTPTTPTTPTTPTTTE 284

RESULT 11
 SGG3_DROER
 ID SGG3_DROER STANDARD; PRT; 328 AA.
 AC P13730;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Salivary glue protein Sgs-3 precursor.
 GN SGG3.
 OS Drosophila erecta (Fruit Fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 CC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OK NCBI_TaxID=7220;
 RN [1]
 RP MEDLINE=883332966; PubMed=3138416;

RA Martin C.H., Mayeda C.A., Meyerowitz E.M.;
 RT "Evolution and expression of the Sgs-3 glue gene of Drosophila.";
 RL J. Mol. Biol. 201:273-287 (1988).
 CC -1- DEVELOPMENTAL STAGE; PRODUCED BY THIRD-INSTAR LARVAE.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M14652; -; NOT ANNOTATED_CDS.
 DR PIR; S01359; S01359.
 DR Flybase; FBgn012268; Dere\SGS3.
 KW Repeat; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 328 SALIVARY GLUE PROTEIN SGS-3.
 SQ SEQUENCE 328 AA; 36355 MW; 62F27F188C0F8272 CRC64;

Query Match 8.8%; Score 179.5; DB 1; Length 328;
 Best Local Similarity 29.9%; Pred. No. 0.00029;
 Matches 67; Conservative 22; Mismatches 122; Indels 13; Gaps 4;

QY 59 PPEMLRNSDTPPLTGPSTPSTVPEPARRSTGLDAGAVTELTTELANMG-----NLS 113
 DB 62 PPTTRRCPPTTRPACCTTKRPTARPTTRRTTVAATTKRATRRRTTKRATRRRTT 121

QY 114 TDSAMEIQTPAATEAQTTPLAATEAQTTRITLA--TEAQTTPLAATEAQTTPPAATEAQT 173
 DB 122 TKRAATRRRTTKRAATRRRTTKRAATRRRTTKRAATRRRTTKRAATRRRTTKRAAT 181

QY 174 TQPTGLEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTP 233
 DB 182 TRRTTVRAITKTKRA---TKRAPTKKATTKRAATPKKATTKRAATPKKATTKRAATPK 237

QY 234 TQPTGLEAQTTPAAMEALSTEPSATREALSMERTTKGLIFPFS 277
 DB 238 TKRAPTKKATTKRAATPKKATTKRAATPKKATTKRAATPKKATTKRAATPKKATTKRA 277

RESULT 12
 VGLG_HSV2H
 ID VGLG_HSV2H STANDARD; PRT; 699 AA.
 AC P13350;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE Glycoprotein G.
 GN GG OR US4.
 OS Herpes simplex virus (type 2 / strain HGS2).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC Alphaherpesvirinae; Simplexvirus.
 OK NCBI_TaxID=10315;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87111457; PubMed=3027242;
 RA McGeech D.J., Moss R.W.M., McNab D., Frame M.C.;
 RT "DNA sequence and genetic content of the HindIII 1 region in the
 RT short unique component of the herpes simplex virus type 2 genome:
 RT identification of the gene encoding glycoprotein G, and evolutionary
 RT comparisons.";
 RL J. Gen. Virol. 68:19-38(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Dolan A.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- MISCELLANEOUS; THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND
 CC 2: GH, GB, GC, GG, GD, GI, AND GE.
 CC -1- MISCELLANEOUS; GLYCOPROTEIN G IS MUCH LARGER IN HSV-2 THAN IN
 CC HSV-1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X04798; CA28490.1; -
 DR EMBL; Z86099; CAB06711.1; -
 DR PIR; C43674; C43674.
 DR InterPro; IPR003363; Herpes_gg.
 DR Pfam; PF02400; Herpes_gg; 1.
 DR KW Glycoprotein.
 FT CARBOHYD 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 512 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 699 AA; 72243 MW; 612BA7B345E75540 CRC64;

Query Match
 Best Local Similarity 8.7%; Score 176.5; DB 1; Length 699;
 Matches 95; Conservative 49; Mismatches 167; Indels 105; Gaps 17;

QY 24 WA----DEAEKALGPIIARDRRO-ATEYEVLDYDF-----LPETPEMLRNST 67
 DB 300 MARGALDDGYPAPPPRRFRFRALRTDDEGVDPDVAAPTRRLMALTEBTSSDSPTSAF 359

QY 68 DTPLTGGPTESTYVPAARRSTGLDAGAVT---ELTTELANMGNLSTDSAAMEIQTT 124
 DB 360 EKPLPVSATAMPSPVPSAEPF---APATYTPPEMNAQAVTAVPEETLV---AS 411

QY 125 QPAAATEAQTPLAATAEQT-----TRULATEAQTPLAATAEQTTPPAATEAQTPTGI 179
 DB 412 PPAATAVESSEPLAAAAAATPGAGHTMTSSSAKATP-PTTPAATTPPTPTSTHATPTPTP 470

QY 180 EAQTT-----APAAEQAQTTPAAEQAQTTPAAEQAQTTPAAEQAQTTPAAEQAQT 232
 DB 471 GPDTTTPPAPTPPVGASAPRTADSPPLTA--SPPAVAPGSAANVAATATPAGTRGTA 528

QY 233 TTOPP-----ATEAQTPLAMEAL 252
 DB 529 RRPPTPKTHPHGPADAPPPSPAPPPENHRGSEFEFGADGPPEDDSATLARTPTN 588

QY 253 STEPSATEALSMPTTKRGLIFPFSVSVTHKGI PMAASNLVNVYVPG-APDHSYKQCL 311
 DB 589 PKNPPAPAPRPPIPTLPPGILGGLAINTPR-----PDAQAP-AKDMSSGPRPHIPLFWPL 643

QY 312 LA-----LIIIALVATTFVCTVLAARL-----SRKGMYPVNNYSPTENVCI 355
 DB 644 TASPALDILFIIISTTHTAFAVCLVALAQLMRGRAGARRRRAHPVRY-----VCL 694

RESULT 13
 CH12_COCCIM STANDARD; PRT; 860 AA.
 AC P54197;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Endochitinase 2 precursor (BC 3.2.1.14).
 OS Coccidioides immitis.
 OC Coccidioides immitis.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Orygenales; mitosporic Orygenales; Coccidioides.
 OX NCBI_TaxID=5501;
 RN RP SEQUENCE FROM N.A.
 RC STRAIN=C735;
 RX MEDLINE=96144270; Pubmed=8566773;
 RA "Isolation and characterization of two chitinase-encoding genes
 RT (cst1, cst2) from the fungus Coccidioides immitis.";

RL Gene 167:173-177(1995).
 CC -1- FUNCTION: MAY BE ASSOCIATED WITH ENDOSPORULATION.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
 CC acetyl-D-glucosamine polymers of chitin.
 CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
 CC HYDROLASES).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; I41662; AAA92642.1; -
 DR HSSP; P23472; 2HW.
 DR InterPro; IPR001579; Chitinase_18/2.
 DR InterPro; IPR001223; Glyco_hydro_18.
 DR Pfam; PF00704; Glyco_hydro_18; 1.
 DR PROSITE; PS01095; CHITINASE_18; 1.
 KW Hydrolyase; Glycosidase; Chitin degradation; Chitin-binding; Signal;
 KW Glycoprotein.
 FT SIGNAL 1 22 POTENTIAL.
 FT CARBOHYD 23 860 ENDOCHITINASE 2.
 FT CHAIN 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 860 AA; 91395 MW; 5E34B54FAA663F3C CRC64;

Query Match
 Best Local Similarity 8.5%; Score 173.5; DB 1; Length 860;
 Matches 82; Conservative 53; Mismatches 141; Indels 93; Gaps 14;

QY 12 LGPGNSLQMPDWADEAEKALGPIIARDRROATEYEVLDYDFLP-----ETEP 59
 DB 288 LPPGENTKIVSYMAKYPSTFGMW---WEATASENKNGGLPAPADIMKEVILACDPPR 344

QY 60 PMLRNST-----DTPLTGGPTEP-TPVPAARRSTGLDAGAVTEL 102
 DB 345 PNSTVSTTSASSTQTSQSTTMTETKTLASSTPSSPSTVAPSSMTQSTSTGSIETV 404

QY 103 TTELAN--MGNLSTDSAAMEIQTTQPAATEAQTTP-----LAATEAQTPL 146
 DB 405 TTRSOEPPSTTSTSA-----STEPVYTRSQEPPSTTSTRSASTEVVTRSQEPPSTT 459

QY 147 -----TATEAQTTP-----LAATEAQTTPPAATEAQTTP 176
 DB 460 ISTWSASTSTSSQDSPTTSTKSAPRTGYVTRSQDLPSTTSTRSEPTETPATTKS 519

QY 177 TGLEAQTTPAAEQAQTTPAAEQAQTTPAAEQAQTTPAAEQAQTTPAAEQAQTTP 236
 DB 520 QCSPIITLSTRSSSAEIVSTRSQHSSSTTISTKSAPTETGTTSEHSTGSPVSTRSASST 579

QY 237 TATEAQTPLAMEALSTEPSATEALSMPTTKRGLIFPFSVSVTHKGI PMAASNLVNV 296
 DB 580 VITRSQNSDSQM-TVSTRSPSTESI-----TTRSQGSSPTFST--KSVPV--DTISTE 629

QY 297 YPVGAPDHI 305
 DB 630 LFSQTPPTI 638

RESULT 14
 ZAN_MOUSE STANDARD; PRT; 5376 AA.
 AC O88789; O08647;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Zonadhesin precursor.
 DN ZAN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OK NCBI_TaxID=10090;
 RN (1)
 RN SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=98123114; PubMed=9452463;
 RA Gao Z., Garbers D.L.;
 RT "Species diversity in the structure of zonadhesin, a sperm-specific
 RT membrane protein containing multiple cell adhesion molecule-like
 RT domains";
 RT J. Biol. Chem. 273:3415-3421 (1998).
 RN [2]
 RP SEQUENCE OF 4864-5376 FROM N.A.
 RC TISSUE=Testis;
 RA MEDLINE=97271566; PubMed=9126492;
 RA Gao Z., Hartup T., Garbers D.L.;
 RT "Chromosome localization of the mouse zonadhesin gene and the human
 RT zonadhesin gene (ZAN)";
 RL Genomics 41:119-122(1997).
 CC -1- FUNCTION: BINDS IN A SPECIES-SPECIFIC MANNER TO THE ZONA PELLUCIDA
 CC OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNITION AND/OR
 CC SIGNALING.
 CC -1- SUBUNIT: PROBABLY FORMS COVALENT OLIGOMERS.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXCLUSIVELY ON THE
 CC APICAL REGION OF THE SPERM HEAD.
 CC -1- TISSUE SPECIFICITY: IN TESTIS; PRIMARILY IN HAPLOID SPERMATIDS.
 CC -1- DOMAIN: THE MAM DOMAINS PROBABLY MEDIATES SPERM ADHESION TO THE
 CC ZONA PELLUCIDA.
 CC -1- DOMAIN: DURING SPERM MIGRATION THROUGH THE REPRODUCTIVE TRACTS,
 CC THE MUCIN-LIKE DOMAIN MIGHT INHIBIT INAPPROPRIATE TRAPPING OF
 CC SPERMATOZOON OR PROMOTING ADHESION TO THE OVIDUCTAL ISTHMUS.
 CC -1- DOMAIN: THE WFMD DOMAIN 2 MAY MEDIATE COVALENT
 CC OLIGOMERIZATION (BY SIMILARITY TO HUMAN INTESTINAL MUCIN MUC2).
 CC -1- SIMILARITY: CONTAINS 3 MAM DOMAINS.
 CC -1- SIMILARITY: CONTAINS 25 WFMD DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/))
 CC -----
 CC EMBL; U97068; AAC63680.1; -
 DR EMBL; U83190; AAC33125.1; -
 DR MGD; MGI:106656; Zan.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR003645; FOLN.
 DR InterPro; IPR000998; MAM domain.
 DR InterPro; IPR002919; TIL_Cystrich.
 DR InterPro; IPR003338; TIL_Cystrich.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF00094; vwd; 4.
 DR Pfam; PF01826; TIL; 25.
 DR Pfam; PF02345; Tila; 25.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00274; FOLN; 11.
 DR SMART; SM00137; MAM; 2.
 DR SMART; SM00214; VMC; 17.
 DR SMART; SM00216; VMD; 4.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 18.
 DR PROSITE; PS00740; MAM_1; FALSE NEG.
 DR PROSITE; PS50060; MAM_2; 3.
 KW Signal; Glycoprotein; Transmembrane; EGF-like domain; Cell adhesion;
 KW Repeat.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 5376 ZONADHESIN.
 FT DOMAIN 18 5317 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 5311 5337 POTENTIAL.

FT	DOMAIN	5338	5376	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	45	210	MAM 1.
FT	DOMAIN	215	374	MAM 2.
FT	DOMAIN	377	542	MAM 3.
FT	DOMAIN	547	1170	80 X HEPTAPEPTIDE REPEATS (APPROXIMATE) (MUCIN-LIKE DOMAIN).
FT	DOMAIN	1171	1280	WFMD 1 (PARTIAL).
FT	DOMAIN	1281	1669	WFMD 2.
FT	DOMAIN	1670	2056	WFMD 3.
FT	DOMAIN	2057	2459	WFMD 4.
FT	DOMAIN	2460	2579	WFMD 5 (PARTIAL).
FT	DOMAIN	2580	2639	WFMD 6 (PARTIAL).
FT	DOMAIN	2700	2819	WFMD 7 (PARTIAL).
FT	DOMAIN	2820	2939	WFMD 8 (PARTIAL).
FT	DOMAIN	2940	3059	WFMD 9 (PARTIAL).
FT	DOMAIN	3060	3179	WFMD 10 (PARTIAL).
FT	DOMAIN	3180	3299	WFMD 11 (PARTIAL).
FT	DOMAIN	3300	3416	WFMD 12 (PARTIAL).
FT	DOMAIN	3417	3536	WFMD 13 (PARTIAL).
FT	DOMAIN	3537	3656	WFMD 14 (PARTIAL).
FT	DOMAIN	3657	3776	WFMD 15 (PARTIAL).
FT	DOMAIN	3777	3892	WFMD 16 (PARTIAL).
FT	DOMAIN	3893	4298	WFMD 17 (PARTIAL).
FT	DOMAIN	4029	4148	WFMD 18 (PARTIAL).
FT	DOMAIN	4149	4263	WFMD 19 (PARTIAL).
FT	DOMAIN	4264	4283	WFMD 20 (PARTIAL).
FT	DOMAIN	4384	4503	WFMD 21 (PARTIAL).
FT	DOMAIN	4504	4623	WFMD 22 (PARTIAL).
FT	DOMAIN	4624	4743	WFMD 23 (PARTIAL).
FT	DOMAIN	4744	4863	WFMD 24 (PARTIAL).
FT	DOMAIN	4864	5261	WFMD 25.
FT	DOMAIN	5259	5295	EGF-LIKE.
FT	DISULFID	5263	5274	BY SIMILARITY.
FT	DISULFID	5268	5293	BY SIMILARITY.
FT	DISULFID	5285	5294	BY SIMILARITY.
FT	CARBOHYD	339	339	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	499	499	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	1216	1216	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	1239	1239	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	1314	1314	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	1814	1814	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	1908	1908	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	1933	1933	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	2028	2028	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	2111	2111	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	2142	2142	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	2332	2332	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	2533	2533	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	2575	2575	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	2692	2692	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	2812	2812	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	3052	3052	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	3065	3065	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	3144	3144	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	3172	3172	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	3288	3288	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	3292	3292	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	3782	3782	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	4005	4005	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	4136	4136	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	4243	4243	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	4254	4254	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	4335	4335	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	4376	4376	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	4586	4586	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	5136	5136	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	5252	5252	N-LINKED (GLCNAc . . .) (POTENTIAL).
SO	SEQUENCE	5376 AA;	579908 MW;	0E44DB77DF2A2620 CRC64;

Query Match 8.4%; DB 1; Length 5376;
 Best Local Similarity 28.6%; Pred. No. 0.018;
 Matches 66; Conservative 24; Mismatches 107; Indels 34; Gaps 9;

QY 57 TEPPEMLRNSTDDTTPPLTGPSTPESTVPEPARRSTGHDAGAVTELTTELANMGNLSTDS 116
 Db 948 TLPTLEVL-----TVPLEVTTPFPGETVTPVEVPTVSTEMVGHTEVTVTPPESTSIPT 1000
 QY 117 AAMEIQTTPA--TEAQTTPLAATEAQTTRLTAT---EAQTTPLA-----TEA 161
 Db 1001 ---EVAIVLVAASLPPEBTTPP---TEVTTTTPEBTTPPAVTVTPPASPPEBTASIVTV 1054
 QY 162 QTPPAATEAQTTPGTEAQTTPAAMEAQTTPAAM--EAQTTPPAAMEAQTTPATM 219
 Db 1055 TTTTPPEBTTPTEVTVTPPEKTIITP--EVTTVPPASIFPEBTTPPEBTTPASEBTTS 1112
 QY 220 EAQTT--APPATEAQTTPPAATEAQTTPLAAMEALSTEPSATLSEPTT 268
 Db 1113 TQTTTLTTEQSAVTOYSIACRPPCPSPPLMIPIGLIPLSKPGVSMFSLAFTT 1163

RESULT 15
 HKR1 YEAST STANDARD; PRT; 1802 AA.

AC P41809;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Hansenula MRAK11 killer toxin-resistant protein 1 precursor.
 GN HKR1 OR YDR420W.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YN 295;
 RX MEDLINE=94156857; Pubmed=8113191;
 RA Kasahara S., Yamada H., Mio T., Shiratori Y., Miyamoto C.,
 RA Yabe I., Nakajima T., Ichishima E., Furuichi Y.,
 RT "Cloning of the Saccharomyces cerevisiae gene whose overexpression
 RT overcomes the effects of HM-1 killer toxin, which inhibits
 RT beta-glucan synthesis.";
 RL J. Bacteriol. 176:1488-1499(1994).
 CC -1- FUNCTION: COULD REGULATE BETA-GLUCAN SYNTHESIS. OVEREXPRESSION
 CC PROVIDES RESISTANCE TO HM-1 KILLER TOXIN.
 CC -1- SUBCELLULAR LOCATION: TYPE I membrane protein (Probable).
 CC -1- PTM: COULD BE O-GLYCOSYLATED IN SERINE/THREONINE RICH DOMAIN.
 CC -1- SIMILARITY: SOME, TO YEAST MSB2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; S69101; AAB30051.1; -.
 DR .SGD; S0002828; HKR1.
 DR KW Glycoprotein; Transmembrane; Repeat; Signal.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 1802 HANSENULA MRAK11 KILLER TOXIN-RESISTANT
 FT PROTEIN 1.
 FT TRANSMEM 1486 1506 POTENTIAL.
 FT DOMAIN 23 1478 SER/THR-RICH.
 FT 453 788 12 X 28 AA TANDEM REPEATS OF S-[AV]-[P]-
 FT V-A-V-S-S-T-Y-T-S-S-P-S-A-P-A-A-I-S-S-T-
 FT Y-T-S-S-P-
 FT 1 (APPROXIMATE).
 FT REPEAT 453 480
 FT REPEAT 481 508
 FT REPEAT 509 536
 FT REPEAT 537 564
 FT REPEAT 565 592
 FT REPEAT 593 620
 FT REPEAT 621 648
 FT REPEAT 649 676

FT REPEAT 677 704 9.
 FT REPEAT 705 732 10.
 FT REPEAT 733 760 11.
 FT REPEAT 761 788 12.
 FT CARBOHYD 24
 FT CARBOHYD 1252 1252 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1293 1293 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1342 1342 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1400 1400 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1802 AA; 188990 MW; E344CA6469785A24 CRC64;

Query Match 8.3%; Score 169.5; DB 1; Length 1802;
 Best Local Similarity 23.7%; Pred. No. 0.0071;
 Matches 86; Conservative 63; Mismatches 145; Indels 69; Gaps 15;

QY 66 SDDTTPPLTGPSTPESTVPEPARRSTGHDAGAVTELTTELANMGNLSTDSAMEIQTTP 125
 Db 478 SSSVPAVAVSSSTYSSPSAPAAISSSTYSSPSAPVAVSSTYT---SSPSAPAAISSTY 532
 QY 126 PATEAQTTPLAATEAQTTRLTAT---TEAQTTPLAATEAQTTPPAATEAQTTPPT 177
 Db 533 ---TSSPSAPVAVSSTYSSPSAPAAISSSTYSSPSAPVAVSSTYT---SSPSAPAAISSTY 589
 QY 178 GLEAQTTPAAMEAQTTP---APAMEAQTTP---PAAAMEAQTTPAAMEAQTTPA 224
 Db 590 ---SSPSAPVAVSSTYSSPSAPAAISSSTYSSPSAPVAVSSTYSSPSAPAAISSTYSS 646
 QY 225 ADE---ATEAQTTPLA-----TEAQTTPLAAMEALSTEPSATLSEPTTAKGL 272
 Db 647 SPSVPAVAVSSTYSSPSAPAAISSSTYSSPSVPAVAVSSTYSSPSAPAAISSTYSSPS 705
 QY 273 FTFPSVSVTHKGIIPMAASNLISVNY-----PV-----GAPDHISVQCLLAIDL 316
 Db 706 -APVAVSS--TYTSSPSAPAAISSSTYSSPSAPVAVSSTYSSPSAPAAISSTYSSPSAP 763
 QY 317 LALVATIFVCTVLAARLSRKGHPVPAVNSPTTEWVCISLLPDGGESPTANGGLSK 376
 Db 764 VAVSSTYSSPSALVLSSTYSSPSYDI-VVSPSTFAAIS-----GYTPSPSASVAVSS 817
 QY 377 AKS 379
 Db 818 TSS 820

Search completed: July 14, 2003, 06:15:13
 Job time : 27 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 14, 2003, 06:10:58 ; Search time 80 Seconds
(without alignments)
1035.387 Million cell updates/sec

Title: US-09-825-580-2

Perfect score: 2030

Sequence: 1 MFLQILLILILGPGNSIQ.....TPPRDRRGGDLTLHSFLP 402

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPPREMBL_21.*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_virus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	824.5	40.6	397	11 Q99L34	Q99L34 mus musculu
2	283	13.9	867	12 Q39782	Q39782 equine herp
3	271.5	13.4	866	12 Q39781	Q39781 equine herp
4	269	13.3	372	12 Q8V0M3	Q8V0M3 equine herp
5	263	13.0	357	12 Q8V0M2	Q8V0M2 equine herp
6	261.5	12.9	342	12 Q8V0L8	Q8V0L8 equine herp
7	261	12.9	826	12 Q8V0L5	Q8V0L5 equine herp
8	259	12.8	337	12 Q8V0M1	Q8V0M1 equine herp
9	256.5	12.6	750	12 Q39307	Q39307 equine herp
10	256	12.5	356	12 Q8V0L7	Q8V0L7 equine herp
11	253.5	12.3	1079	5 Q9N4S7	Q9N4S7 caenorhabdi
12	249.5	12.0	293	12 Q8V0K1	Q8V0K1 equine herp
13	244	12.0	316	12 Q8V0M4	Q8V0M4 equine herp
14	243.5	12.0	316	12 Q8V0M0	Q8V0M0 equine herp
15	241	11.9	389	12 Q8V0L6	Q8V0L6 equine herp
16	239.5	11.8	374	12 Q8V0L6	Q8V0L6 equine herp

RESULT 1 ID	Q99L34	PRELIMINARY;	PRT;	397 AA.	
AC	01-JUN-2001	(TRMBLrel. 17, Created)			
DT	01-JUN-2001	(TRMBLrel. 17, Last sequence update)			
DT	01-JUN-2002	(TRMBLrel. 21, Last annotation update)			
DE	Similar to selectin, platelet (p-selectin) ligand.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				Q8V0L3 equine herp
CC	Mammalia; Eutheria; Rodentia; Sclirognathi; Muridae; Murinae; Mus.				Q8V0L2 equine herp
OX	NCBI_TaxID=10090;				Q8V0K0 equine herp
RN	(1)				Q23635 caenorhabdi
RP	SEQUENCE FROM N.A.				Q96W8 schizosacch
RA	Strausberg R.;				Q8V0L1 equine herp
RL	Submitted (FEH-2001) to the EMBL/GenBank/DBJ databases.				Q8V0L0 equine herp
DR	EMBL; BC003874; AH03874.1; ..				Q8V0K8 equine herp
DR	MGD; MGI:106689; Selpl.				Q8V0L4 equine herp
KW	Lectin; Selectin.				Q8V0K5 equine herp
SQ	SEQUENCE 397 AA; 41844 MW; 552C7791622D23A3 CRC64;				Q17921 caenorhabdi

ALIGNMENTS

Query Match	Best Local Similarity	Score	824.5	DB 11;	Length	397;			
Matches	193;	Conservative	53;	Mismatches	119;	Indels	69;	Gaps	8;
Qy	1	MFLQILLILILGPGNSIQ...TPPRDRRGGDLTLHSFLP	59						
Db	1	MSPPLVILITIGPNSIQ...DPPWHFKKANGPVLHERRQVVGDDDDPDDYVNTDP	60						
Qy	60	PMLRNSTDTPL.....TGPTPSTVEPARRSGLDAGAVTELT	104						
Db	61	PELLKAVNTVAHDELPTVWMLRDSAGTSRAVEKIKITDPTAAGTGT	115						
Qy	105	ELANNGNSTDSAAEIQTPAATAEQTPLAATAEQTTLTATEAQTPLAATAEQT	164						
Db	116	---VGLSTDSA-----TQMSLTSVETVQPASTVEVTS	145						
Qy	165	PPAATAEQTOPGLAEQTAAPAAEQAQTAPAAEQAQTTPAAEQAQTTPAAEQAQT	224						

Query Match 13.3%; Score 269; DB 12; Length 372;
 Best Local Similarity 37.5%; Pred. No. 7.2e-09;
 Matches 78; Conservative 9; Mismatches 115; Indels 6; Gaps 1;

QY 67 TDTPTTGTGCTPESTTVEPARRSGLDAGAVTELTTELANKNLSSTSSAMEIQTQP 126
 DB 158 TTTTPTSTTTTAAVTTTAAVTTTAAVTTTAAVTTTAAVTTTAAVTTTAAVTTTAA 217
 QY 127 AATEAQTPLAATEAQTPLAATEAQTPLAATEAQTPLAATEAQTPLAATEAQTPLA 186
 DB 218 ATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTA 277
 QY 187 AAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTP 246
 DB 278 ATSSAHTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 337
 QY 247 AAMEALST----EPSATEALSMETT 268
 DB 338 PFGSSTSTTGASTSTPSASTATSATPTS 365

RESULT 5
 Q8V0M2 PRELIMINARY; PRT; 357 AA.
 AC Q8V0M2; (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Glycoprotein gp2 (Fragment).
 OS Equine herpesvirus 1.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10326;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1026A/93;
 RA Huang J.-A., Ficorilli N., Hartley C.A., Allen G.P., Studdert M.J.;
 RT "Polymorphism of open reading frame 71 of equine herpesvirus 4 (EHV4)
 RT and EHV1."
 RL Submitted (MAY-2001) to the EMBL/Genbank/DBD databases.
 DR EMBL; AY034638; AAK61478.1; --
 DR InterPro; IPR000104; Antifreeze_1.
 DR PRINTS; PR00308; ANTI-FREZEI.
 FT NON TER 357 357
 SQ SEQUENCE 357 AA; 33054 MW; 08310EF29F5B2641 CRC64;

Query Match 13.0%; Score 263; DB 12; Length 357;
 Best Local Similarity 33.3%; Pred. No. 1.6e-08;
 Matches 83; Conservative 11; Mismatches 139; Indels 16; Gaps 3;

QY 59 PPEMLNSDTPPLPGPPESTTVEPARRSGLDAGAVTELTTELANKNLSSTSSAMEIQTQP 118
 DB 101 PFTSTETTTTTPASTTPTTPTTAAVTTTAAVTTTAAVTTTAAVTTTAAVTTTAAVTTT 160
 QY 119 MEIQTQPA-----TEAQTPLAATEAQTPLAATEAQTPLAATEAQTPLAATEAQT 173
 DB 161 TPTSTTTTATTTVPTASTTDTTAAATTTAATTTAATTTAATTTAATTTAATTTAAT 220
 QY 174 TQFTGLEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTP 233
 DB 221 TAATTTAAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 279
 QY 234 TQPTATEAQTPLAAMEALSTPSATEALSMETTTRKGLFIPSSVYHKGIPLMAASND 293
 DB 280 TAAATTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 329
 QY 294 SVNYPVGPAP 302
 DB 330 TTGASTSTP 338
 RESULT 6

Q8V0L8 PRELIMINARY; PRT; 342 AA.
 AC Q8V0L8; (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Glycoprotein gp2 (Fragment).
 OS Equine herpesvirus 1.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10326;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1077/94;
 RA Huang J.-A., Ficorilli N., Hartley C.A., Allen G.P., Studdert M.J.;
 RT "Polymorphism of open reading frame 71 of equine herpesvirus 4 (EHV4)
 RT and EHV1."
 RL Submitted (MAY-2001) to the EMBL/Genbank/DBD databases.
 DR EMBL; AY034642; AAK61482.1; --
 DR InterPro; IPR000104; Antifreeze_1.
 DR PRINTS; PR00308; ANTI-FREZEI.
 FT NON TER 342 342
 SQ SEQUENCE 342 AA; 31718 MW; 3A0D3BE0DEF0D688 CRC64;

Query Match 12.9%; Score 261.5; DB 12; Length 342;
 Best Local Similarity 34.9%; Pred. No. 1.8e-08;
 Matches 82; Conservative 10; Mismatches 118; Indels 25; Gaps 3;

QY 59 PPEMLNSDTPPLPGPPESTTVEP--AARRSGLDAGAVTELTTELANKNLSST-- 114
 DB 101 PFTSTETTTTTPASTTPTTPTTAAVTTTAAVTTTAAVTTTAAVTTTAAVTTTAAVTTT 160
 QY 115 -----DSAMEIQTQPAATEAQTPLAATEAQTPLAATEAQTPLAATEAQTPLAAT 159
 DB 161 TPTSTTTTATTTVPTASTTDTTAAATTTAATTTAATTTAATTTAATTTAATTTAAT 220
 QY 160 EAQTTPAATEAQTTPGLEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQT 219
 DB 221 TAATTTAAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 280
 QY 220 EAQTTPAATEAQTTPAATEAQTTPAAMEALST-----EPSATEALSMETT 268
 DB 281 AATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 335

RESULT 7
 Q8V0L5 PRELIMINARY; PRT; 826 AA.
 AC Q8V0L5; (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Glycoprotein gp2.
 OS Equine herpesvirus 1.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10326;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=438/77;
 RA Huang J.-A., Ficorilli N., Hartley C.A., Allen G.P., Studdert M.J.;
 RT "Polymorphism of open reading frame 71 of equine herpesvirus 4 (EHV4)
 RT and EHV1."
 RL Submitted (MAY-2001) to the EMBL/Genbank/DBD databases.
 DR EMBL; AY034645; AAK61485.1; --
 DR InterPro; IPR000104; Antifreeze_1.
 DR PRINTS; PR00308; ANTI-FREZEI.
 FT NON TER 826 826
 SQ SEQUENCE 826 AA; 82915 MW; F5B61157739CA74D CRC64;

Query Match 12.9%; Score 261; DB 12; Length 826;
 Best Local Similarity 29.4%; Pred. No. 5e-06;
 Matches 102; Conservative 25; Mismatches 180; Indels 40; Gaps 9;

Db 767 T-----TTAETSTTEPSSSNTPV 786

RESULT 13

Q8V0K1 PRELIMINARY; PRT; 293 AA.

AC Q8V0K1; PRELIMINARY; PRT; 293 AA.
01-MAR-2002 (TEMBLrel. 20, Created)
01-MAR-2002 (TEMBLrel. 20, Last sequence update)
01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Glycoprotein gp2 (Fragment).
OS Equine herpesvirus 4.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicelloviruses.
NCBI_TaxID=10331;

RP SEQUENCE FROM N.A.
RC STRAIN=R19/68;
RA Huang J.-A., Ficorilli N., Hartley C.A., Allen G.P., Studdert M.J.;
RT "Polymorphism of open reading frame 71 of equine herpesvirus 4 (EHV4)
RL and EHV1."
RL Submitted (May-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY034660; AAKS8439.1;
DR InterPro; IPR000104; Antifreeze_1.
DR PRINTS; PR00308; ANTIFREEZEI.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 293 AA; 27805 MW; 7432PEA4D5BE9PF CRC64;

Query Match 12.0%; Score 244; DB 12; Length 293;
Best Local Similarity 37.4%; Pred. No. 1.8e-07;
Matches 80; Conservative 11; Mismatches 111; Indels 12; Gaps 4;

QY 57 TEPPMRLNSTDTTPLTGPSTVEPARRSTGIDAGAVTELTTELANMGNLSTDS 116
3 TSVSEPTSTTAAAT-PTTEST--AATTTSTTAAATTTTSTSTSTSTA 57
117 AAMEIOTTQPAATEAOTTPAATEAOTTRLTATEAOTTPAATEAOTTPAATEAOTTP 176
58 ATTTTSTTAAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTA 115
177 TGEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTP 236
116 --TAAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTA 173
237 TATEAOT--TPLAAMENLSTEPSATELSMEPT 267
174 TTTAATTTAATPTSESEAVSTLAATTAADTT 207

RESULT 14
Q8V0M4 PRELIMINARY; PRT; 316 AA.

AC Q8V0M4; PRELIMINARY; PRT; 316 AA.
01-MAR-2002 (TEMBLrel. 20, Created)
01-MAR-2002 (TEMBLrel. 20, Last sequence update)
01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Glycoprotein gp2 (Fragment).
OS Equine herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicelloviruses.
NCBI_TaxID=10326;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=848/89;
RA Huang J.-A., Ficorilli N., Hartley C.A., Allen G.P., Studdert M.J.;
RT "Polymorphism of open reading frame 71 of equine herpesvirus 4 (EHV4)
RL and EHV1."
RL Submitted (May-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY034636; AAK61476.1;
DR InterPro; IPR000104; Antifreeze_1.
DR PRINTS; PR00308; ANTIFREEZEI.
FT NON_TER 316

SQ SEQUENCE 316 AA; 29361 MW; F8443E6275E06DA8 CRC64;

Query Match 12.0%; Score 243.5; DB 12; Length 316;
Best Local Similarity 35.5%; Pred. No. 2e-07;
Matches 75; Conservative 13; Mismatches 120; Indels 3; Gaps 3;

QY 59 PEMLRNSTDTTPLTGPSTVEPARRSTGIDAGAVTELTTELANMGNLSTDSAA 118
101 PNSTAETTTTPTASTTTPTTTTAAATTAATTAATTAATTAATTAATTAATTAAT 160
119 MEIOTTQPAATEAOTTPAATEAOTTRLTATEAOTTPAATEAOTTPAATEAOTTP 178
161 TPTSTTTTAAATTTVPT-TASTTDTTAAATTTAATTTAATTTAATTTAATTTAAT 219
179 LEAQTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPA 238
220 TTRATTS-SATTAATTSATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 278
239 TEAOTTPAAMEALSTE-PSATEALSMEPTT 268
279 TGSPTSGSTTGAATSTPSASATSAATPTS 309

RESULT 15

Q8V0M0 PRELIMINARY; PRT; 389 AA.

AC Q8V0M0; PRELIMINARY; PRT; 389 AA.
01-MAR-2002 (TEMBLrel. 20, Created)
01-MAR-2002 (TEMBLrel. 20, Last sequence update)
01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Glycoprotein gp2 (Fragment).
OS Equine herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicelloviruses.
NCBI_TaxID=10326;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1070/94;
RA Huang J.-A., Ficorilli N., Hartley C.A., Allen G.P., Studdert M.J.;
RT "Polymorphism of open reading frame 71 of equine herpesvirus 4 (EHV4)
RL and EHV1."
RL Submitted (May-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY034640; AAK61480.1;
DR InterPro; IPR000104; Antifreeze_1.
DR PRINTS; PR00308; ANTIFREEZEI.
FT NON_TER 389
FT NON_TER 389
SQ SEQUENCE 389 AA; 42357630A8C3A592 CRC64;

Query Match 11.9%; Score 241; DB 12; Length 389;
Best Local Similarity 35.7%; Pred. No. 3.6e-07;
Matches 82; Conservative 14; Mismatches 114; Indels 20; Gaps 5;

QY 55 PETEPMRLNSTDTTPLTGPSTVEPARRSTGIDAGAVTELTTELANMGNLST 114
157 PTTTTPTSTTTTAAAT--TVPTTASSTTDTTAAATTTAATTTAATTTAATTTAAT 214
115 DSAAMEIOTTQPAATEAOTTPAATEAOTTRLT-----ATEAOTTPA-----AT 159
215 TTAATTTAATTSATTAATTSATTTAATTTAATTTAATTTAATTTAATTTAATTSAT 274
160 EAQTTPAATEAOTTPAATEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPA 219
275 TAATTSATTAATTSAT--TAATTSATTAATTTAATTTAATTTAATTTAATTTAAT 332
220 EAQTTPAATEAOTTPAATEAOTTPAAMEALSTE-PSATEALSMEPTT 268
333 AATTTAATTTAATTTAATTTAATTTGSPGTSSTGASTSPTSAATSAATPTS 382

Search completed: July 14, 2003, 06:14:41
Job time : 83 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comogen Ltd.

OM protein - protein search, using sw model

Run on: July 14, 2003, 06:10:56 ; Search time 72 Seconds
(without alignments)
743.983 Million cell updates/sec

Title: US-09-825-580-2

Perfect score: 2030
Sequence: 1 MFLQLLLILLGPGNSIQI.....TPPRDRRGGDDLTTHSFUP 402

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: A_Geneseq_101002.*
 - 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
 - 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
 - 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
 - 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
 - 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
 - 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
 - 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
 - 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
 - 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
 - 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
 - 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
 - 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
 - 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
 - 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
 - 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
 - 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
 - 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
 - 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
 - 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
 - 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
 - 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
 - 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARYS

Result No.	Score	Query Match	Length	ID	Description
1	2030	100.0	402	AA85075	P-selectin ligand.
2	2030	100.0	402	AAW26174	P-selectin glycoprotein
3	2030	100.0	402	AAW72695	Human glycoprotein
4	2030	100.0	402	AAW53320	P-selectin ligand
5	2030	100.0	402	AAW29765	Human P-selectin 1
6	2030	100.0	402	AAW81966	Human P-selectin 1
7	2030	100.0	402	AAW81011	Human P-selectin 1
8	2030	100.0	402	AAU09941	Human P-selectin g
9	2030	100.0	402	AAU79653	Human P-selectin g
10	2030	100.0	402	AAW7945	Human P-selectin g

Result No.	Score	Query Match	Length	ID	Description
11	2022	99.6	402	AAW53965	P-selectin ligand.
12	2015	99.3	412	AAW72696	Human glycoprotein
13	2015	99.3	412	AAW56346	P-selectin ligand
14	2015	99.3	412	AAW29778	Human P-selectin 1
15	2015	99.3	412	AAW81967	Human P-selectin 1
16	2015	99.3	412	AAW83012	Human placenta P-s
17	2015	99.3	412	ABB04668	PSGL-1 amino acid
18	2011	99.1	412	AAW53966	P-selectin ligand.
19	2004	98.7	412	AAW85076	P-selectin ligand.
20	1503	74.0	425	AAW29774	P-selectin ligand
21	1494	73.5	646	AAW29775	P-selectin ligand
22	1476	72.7	647	AAW29777	P-selectin ligand
23	1475	72.7	528	AAU79654	Human soluble PSGL
24	953.5	47.0	312	AAW29772	P-selectin ligand
25	953	46.9	320	AAW29773	P-selectin ligand
26	507.5	25.0	100	AAW10530	P-selectin ligand
27	475	23.4	438	AAW29776	P-selectin ligand
28	467	23.0	422	AAW29771	P-selectin ligand
29	465	22.9	264	AAW53324	P-selectin ligand
30	465	22.9	264	AAW29769	Human P-selectin 1
31	465	22.9	313	AAW53321	P-selectin ligand
32	465	22.9	313	AAW29766	P-selectin ligand
33	465	22.9	269	AAW7946	Human dimeric P-se
34	463	22.8	269	AAW53322	P-selectin ligand
35	463	22.8	269	AAW29767	Human P-selectin 1
36	458	22.6	201	AAW29770	P-selectin ligand
37	330	16.3	63	AAW26188	Fragment #14 of P-
38	325	16.0	62	AAW26187	Fragment #13 of P-
39	311	15.3	59	AAW26186	Fragment #12 of P-
40	288.5	14.2	307	ABB65879	Drosophila melanog
41	278	13.7	437	AAW53323	P-selectin ligand
42	278	13.7	437	AAW29768	Human P-selectin 1
43	266	13.1	50	AAW26185	Fragment #11 of P-
44	261	12.9	49	AAW26184	Fragment #10 of P-
45	223	11.0	41	AAW26183	Fragment #9 of P-s

ALIGNMENTS

RESULT 1

Result No.	Score	Query Match	Length	ID	Description
AA85075				AA85075	standard; Protein; 402 AA.
XX				XX	
AC				AA85075;	
XX				XX	
DT				01-MAR-1996 (first entry)	
XX				XX	
DE				P-selectin ligand.	
XX				XX	
KW				P-selectin ligand; adhesion; cell adhesion; antiinflammatory.	
XX				XX	
OS				Homo sapiens.	
XX				XX	
FH				Key	Location/Qualifiers
FT				Peptide	1..21
FT				Domain	/label= Sig_peptide
FT				Domain	21..310
FT				Cleavage-site	/label= Extracellular_domain
FT				Modified-site	/label= PACE_cleavage_site
FT				Modified-site	46
FT				Modified-site	/label= Tyr_sulfation_site
FT				Modified-site	48
FT				Modified-site	/label= Tyr_sulfation_site
FT				Modified-site	51
FT				Modified-site	/label= Tyr_sulfation_site
FT				Modified-site	65..67
FT				Modified-site	/label= N-glycosylation_site
FT				Modified-site	111..113
FT				Modified-site	/label= N-glycosylation_site
FT				Modified-site	292..294
FT				Modified-site	/label= N-glycosylation_site

```

FT Region 311..332
FT /Label= Transmembrane_region
FT Domain 333..402
FT /Label= Cytoplasmic_domain
XX
PN MO9530001-A2.
XX
PD 09-NOV-1995.
XX
PF 24-APR-1995; 95WO-US04968.
XX
PR 30-SEP-1994; 94US-0316305.
PR 28-APR-1994; 94US-0235398.
XX
PA (GEMV ) GENETICS INST INC.
XX
PI Chang X, Cumming D, Kumar R, Larsen GR, Sako DS;
PI Shaw GD, Veldman GM;
XX
DR WPI; 1995-403865/51.
DR N-PSDB; AAT02484.
XX
PT Novel P-selectin ligand protein - useful as anti-inflammatory agent
PT and to identify inhibitors of selectin-mediated intercellular
PT adhesion
XX
PS Claim 1; Page 82-83; 135pp; English.
XX
CC P-selectin ligand (AAR85075) is an inhibitor of E- and P-selectin-
CC mediated intercellular adhesion. It was obtd. by expression
CC of cDNA clone PWT21:P85 (ATCC 69096) (AAT02484), isolated from human
CC promyelocytic HL-60 cells. Recombinant P-selectin ligand can
CC be produced in mammalian host cells co-transfected with a
CC fucosyltransferase gene, for use as an antiinflammatory.
XX
SQ Sequence 402 AA;

```

```

XX AC AAW26174;
XX XX 28-NOV-1997 (first entry)
DT 28-NOV-1997
XX DE P-selectin glycoprotein ligand-1.
XX
XX KW P-selectin glycoprotein ligand-1; PSGL-1; human; mucin-like protein;
XX KW calcium-dependent carbohydrate binding protein; platelet; endothelium;
XX KW thrombin; leukocyte; sulphated glycosylated peptide; O-glycan; therapy;
XX KW inflammation; ischaemia; reperfusion; bacterial sepsis; atherosclerosis;
XX KW disseminated intravascular coagulation; rheumatoid arthritis; antibody;
XX KW adult respiratory distress syndrome; tumour metastasis;
XX KW autoimmune disorder; inflammatory disorder.
XX OS Homo sapiens.
XX
XX EH Key Location/Qualifiers
XX FT Peptide 1..19
XX FT Peptide /note= "signal peptide"
XX FT Peptide 19..41
XX FT Peptide /note= "propeptide"
XX FT Protein 42..402
XX FT Protein /note= "mature PSGL-1"
XX FT Domain 42..308
XX FT Domain /note= "extracellular domain"
XX FT Domain 309..333
XX FT Domain /note= "transmembrane domain"
XX FT Domain 334..402
XX FT Domain /note= "cytoplasmic tail"
XX
XX PN MO9706176-A2.
XX
XX PD 20-FEB-1997.
XX
XX PF 02-AUG-1996; 96WO-US12820.
XX
XX PR 17-MAY-1996; 96US-0649802.
XX PR 03-AUG-1995; 95US-0510920.
XX PR 15-MAY-1996; 96US-0017794.
XX
XX PA (OKLA ) UNIV OKLAHOMA STATE.
XX
XX PI Cummings RD, Mcever RP, Moore KL;
XX
XX DR WPI; 1997-154206/14.
XX
XX PT New O-glycan (8) derived from P-selectin glyco-protein ligand
XX PT (PSGL)-1 - used for inhibiting binding of PSGL-1 to selectins, e.g.
XX PT for treating inflammatory or autoimmune disorders or tumours
XX
XX PS Claim 16; Page 78-79; 96pp; English.
XX
XX CC This sequence represents the human P-selectin glycoprotein ligand-1
XX CC (PSGL-1). P-selectin is a calcium-dependent carbohydrate binding protein
XX CC expressed on the surfaces of activated platelets and endothelium in
XX CC response to thrombin and other agonists. PSGL-1 is a high affinity
XX CC P-selectin ligand produced by leukocytes. Binding of P-selectin to
XX CC PSGL-1 is calcium ion dependent and is abolished by treatment of the
XX CC ligand with statilase. PSGL-1 is a homodimer, with a highly extended
XX CC extracellular domain, which is a feature of mucin-like proteins. This
XX CC sequence is targeted by the sulphated glycosylated peptides of the
XX CC invention (see AAW26175-W26188). PSGL-1 is also targeted by the
XX CC O-glycans, and O-glycanated products of the invention. The products are
XX CC used for inhibiting the binding of P-selectin and other selectins to
XX CC PSGL-1. They can be used for the treatment of inflammation, injury
XX CC resulting from ischaemia and reperfusion, bacterial sepsis and
XX CC disseminated intravascular coagulation, adult respiratory distress
XX CC syndrome, tumour metastasis, rheumatoid arthritis, atherosclerosis and
XX CC other autoimmune or inflammatory disorders. The products can also be used
XX CC in the production of antibodies and in diagnostic applications.
XX
XX SQ Sequence 402 AA;

```


Query Match 100.0%; Score 2030; DB 18; Length 402;
Best Local Similarity 100.0%; Pred. No. 2.3e-149;
Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFLQILLLLLLILGPNLSIQMPTWADDEAKKALGPIIARPRQATREYIDYDFLPETERP 60
DB 1 MFLQILLLLLLILGPNLSIQMPTWADDEAKKALGPIIARPRQATREYIDYDFLPETERP 60
QY 61 EMLRNSDTTPTLTGPTPESTTVEPARRSSTGLDAGAVTELTTELANNGNLSTDSAAE 120
DB 61 EMLRNSDTTPTLTGPTPESTTVEPARRSSTGLDAGAVTELTTELANNGNLSTDSAAE 120
QY 121 IQTTQPAATEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPGTE 180
DB 121 IQTTQPAATEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPGTE 180
QY 181 AQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAATE 240
DB 181 AQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAATE 240
QY 241 AQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAATE 300
DB 241 AQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAATE 300
QY 301 APDHISVQCLLAILLALVATTFVCTVVLAVRLSRKGMYPVANNYSPTMWCISSLP 360
DB 301 APDHISVQCLLAILLALVATTFVCTVVLAVRLSRKGMYPVANNYSPTMWCISSLP 360
QY 361 DGGEGPSATANGSLSKAKSPGILTPPRRREGDDDLTHSFLP 402
DB 361 DGGEGPSATANGSLSKAKSPGILTPPRRREGDDDLTHSFLP 402

RESULT 3
AAW72695
ID AAW72695 standard; Protein; 402 AA.

AAW72695;
08-JAN-1999 (first entry)
Human glycoprotein P-selectin ligand protein.
Human; glycoprotein; P-selectin ligand protein; inflammatory disease;
intracellular domain; E-selectin-mediated intercellular adhesion.
Homo sapiens.
Key Location/Qualifiers
FH Misc-difference 42..60
FT /note= "see claim 2"
FT Misc-difference 42..88
FT /note= "see claim 3"
FT Misc-difference 42..118
FT /note= "see claim 5"
FT Misc-difference 42..189
FT /note= "see claim 6"
FT Misc-difference 42..310
FT /note= "see claim 8"
FT Misc-difference 42..316
FT /note= "see claim 10"
US5827817-A.
27-OCT-1998.
07-JUN-1995; 95US-0477254.
25-APR-1995; 95US-0428734.
23-OCT-1992; 92US-0963662.
26-AUG-1993; 93US-0112608.
22-OCT-1993; 93WO-US10168.

PR 28-APR-1994; 94US-0235398.
PR 30-SEP-1994; 94US-0316305.
PR 07-JUN-1995; 95US-0477254.
XX
XX (GENY) GENETICS INSTR INC.
XX
XX Camphausen R, Chang X, Cumming D, Kumar R, Larsen GR;
PI Sako DS, Shaw G, Veldman GM;
DR MPI: 1998-594030/50.
DR N-PSDB; AAW7130.
PT Soluble P-selectin ligand glycoprotein - with truncated protein
PT sequence lacking intracellular domain
PS Claim 1; Column 47-50; 66pp; English.
XX
XX The present sequence represents a human glycoprotein having P-selectin
CC ligand (PSL) activity. The present invention describes a glycoprotein
CC having PSL activity which comprises a carbohydrate and a protein. The
CC protein includes amino acids 42-60 of AAW72695, but lacks the
CC intracellular domain and optionally the transmembrane domain of mature
CC PSL protein, and the carbohydrate contains sialyl Lewis X. The isolated
CC PSL protein may be useful in treating conditions characterized by P- or
CC E-selectin mediated intercellular adhesion. Such conditions include
CC myocardial infarction, bacterial or viral infection, metastatic
CC conditions, inflammatory disorders, asthma, emphysema, thermal injury,
CC multiple sclerosis, diabetes, glomerulonephritis, gingivitis,
CC dermatosis, Grave's disease, hemolytic uremic syndrome, ulcerative colitis, Crohn's
CC peritonitis, necrotizing enterocolitis, and cytokine-induced toxicity. The
CC isolated PSL protein may also be useful in organ transplantation, both to
CC prepare organs for transplantation and to quell organ transplant
CC rejection. The isolated PSL may be used to treat haemodialysis and
CC leukophoresis patients. Additionally, isolated PSL protein may be used as
CC an antimetastatic agent. The isolated PSL protein may be used itself as
CC an inhibitor of P- or E-selectin-mediated intercellular adhesion or to
CC design inhibitors of P- or E-selectin-mediated intercellular adhesion.
XX
XX
SQ Sequence 402 AA;

Query Match 100.0%; Score 2030; DB 19; Length 402;
Best Local Similarity 100.0%; Pred. No. 2.3e-149;
Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MFLQILLLLLLILGPNLSIQMPTWADDEAKKALGPIIARPRQATREYIDYDFLPETERP 60
DB 1 MFLQILLLLLLILGPNLSIQMPTWADDEAKKALGPIIARPRQATREYIDYDFLPETERP 60
QY 61 EMLRNSDTTPTLTGPTPESTTVEPARRSSTGLDAGAVTELTTELANNGNLSTDSAAE 120
DB 61 EMLRNSDTTPTLTGPTPESTTVEPARRSSTGLDAGAVTELTTELANNGNLSTDSAAE 120
QY 121 IQTTQPAATEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPGTE 180
DB 121 IQTTQPAATEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPGTE 180
QY 181 AQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAATE 240
DB 181 AQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAATE 240
QY 241 AQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAATE 300
DB 241 AQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAATE 300
QY 301 APDHISVQCLLAILLALVATTFVCTVVLAVRLSRKGMYPVANNYSPTMWCISSLP 360
DB 301 APDHISVQCLLAILLALVATTFVCTVVLAVRLSRKGMYPVANNYSPTMWCISSLP 360
QY 361 DGGEGPSATANGSLSKAKSPGILTPPRRREGDDDLTHSFLP 402
DB 361 DGGEGPSATANGSLSKAKSPGILTPPRRREGDDDLTHSFLP 402

Db 121 IOTTQPAATEAQTTPLAATEAQTTRLTVAATEAQTTPLAATEAQTTPPAATEAQTTPQTGLE 180
 Qy 181 AQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAATEAQTTPATE 240
 Db 181 AQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAATEAQTTPATE 240
 Qy 241 AQTTPLAAMEALSTEPSATEALSMEPTTKRGLFIPFSVSVTHKGI PMAASNLVNYPVG 300
 Db 241 AQTTPLAAMEALSTEPSATEALSMEPTTKRGLFIPFSVSVTHKGI PMAASNLVNYPVG 300
 Qy 301 APDHISVKOCLLAILLALVAVTTFVCTVVAVLRSLRKGMYPVARNYSPTMWCISISLLP 360
 Db 301 APDHISVKOCLLAILLALVAVTTFVCTVVAVLRSLRKGMYPVARNYSPTMWCISISLLP 360
 Qy 361 DGGEGSATANGLSKAKSPGLTPPEPREREGDDLTTHSFLP 402
 Db 361 DGGEGSATANGLSKAKSPGLTPPEPREREGDDLTTHSFLP 402

RESULT 7
 AAW83011 standard; Protein: 402 AA.

AAW83011;
 28-JAN-1999 (first entry)
 Human promyelocyte P-selectin ligand glycoprotein.

Human, promyelocyte; placenta; P-selectin ligand glycoprotein;
 inflammatory disease; P-selectin-mediated intercellular adhesion;
 E-selectin-mediated intercellular adhesion.
 Homo sapiens.

US5840679-A.
 24-NOV-1998.
 07-JUN-1995; 95US-0472576.
 25-APR-1995; 95US-0428734.
 23-OCT-1992; 92US-0965662.
 26-AUG-1993; 93US-0112608.
 22-OCT-1993; 93MO-US10168.
 28-APR-1994; 94US-0235398.
 30-SEP-1994; 94US-0316305.
 07-JUN-1995; 95US-0472576.
 (GEMV) GENETICS INST INC.

Gamphausen R, Chang X, Gunning D, Kumar R, Larsen GR;
 Sako DS, Shaw G, Veldman GM;
 WPI; 1999-034073/03.
 N-PSDB; AAV69835.

Inhibition of P-selectin-mediated intercellular binding - by
 administering soluble P-selectin ligand glycoprotein
 Claim 1; Column 47-50; 66pp; English.

A method has been developed for inhibiting P-selectin-mediated binding.
 The method comprises administering a soluble P-selectin ligand
 glycoprotein which lacks the transmembrane or intracellular domain of
 the mature glycoprotein and in which the protein portion comprises at
 least amino acids 42-60 i.e. Gln Ala Thr Glu Tyr Leu Asp Tyr
 Asp Phe Leu Pro Glu Thr Glu Pro Pro, and the carbohydrate portion
 comprises sialyl Lewis X or sialyl Lewis A. Isolated P-selectin ligand
 protein may be useful in treating conditions characterised by P- or
 E-selectin mediated intercellular adhesion. Such conditions include
 myocardial infarction, bacterial or viral infection, metastatic
 conditions, inflammatory disorders, thermal injury, multiple sclerosis,

CC diabetes, Reynaud's syndrome, neutrophilic dermatosis, inflammatory
 CC bowel disease, Grave's disease, peridontitis, ulcerative colitis,
 CC Crohn's disease, granulocyte transfusion associated syndrome, and
 CC cytokine-induced toxicity. Isolated P-selectin ligand protein may also
 CC be useful in organ transplantation, both to prepare organs for
 CC transplantation and to quell organ transplant rejection. Isolated
 CC P-selectin ligand protein may be used to treat hemodialysis and
 CC leukophoretic patients. The present sequence represents P-selectin
 CC ligand glycoprotein isolated from human promyelocytes.

Query Match 100.0%; Score 2030; DB 20; Length 402;
 Best Local Similarity 100.0%; Pred. No. 2.3e-149;
 Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPEQLLLILLGPGNSLQIMPTWADAEKALGPIIARRRQATEYVLDYDFLPETEP 60
 Db 1 MPEQLLLILLGPGNSLQIMPTWADAEKALGPIIARRRQATEYVLDYDFLPETEP 60
 Qy 61 EMLRNSTDTTPTGPGTPEPTTVEPARRSTGIDAGAVTELTTELANNGLSTDSAAE 120
 Db 61 EMLRNSTDTTPTGPGTPEPTTVEPARRSTGIDAGAVTELTTELANNGLSTDSAAE 120
 Qy 121 IOTTQPAATEAQTTPLAATEAQTTRLTVAATEAQTTPLAATEAQTTPAATEAQTTPQTGLE 180
 Db 121 IOTTQPAATEAQTTPLAATEAQTTRLTVAATEAQTTPLAATEAQTTPAATEAQTTPQTGLE 180
 Qy 181 AQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAATEAQTTPATE 240
 Db 181 AQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAATEAQTTPATE 240
 Qy 241 AQTTPLAAMEALSTEPSATEALSMEPTTKRGLFIPFSVSVTHKGI PMAASNLVNYPVG 300
 Db 241 AQTTPLAAMEALSTEPSATEALSMEPTTKRGLFIPFSVSVTHKGI PMAASNLVNYPVG 300
 Qy 301 APDHISVKOCLLAILLALVAVTTFVCTVVAVLRSLRKGMYPVARNYSPTMWCISISLLP 360
 Db 301 APDHISVKOCLLAILLALVAVTTFVCTVVAVLRSLRKGMYPVARNYSPTMWCISISLLP 360
 Qy 361 DGGEGSATANGLSKAKSPGLTPPEPREREGDDLTTHSFLP 402
 Db 361 DGGEGSATANGLSKAKSPGLTPPEPREREGDDLTTHSFLP 402

RESULT 8
 AAU09941
 ID AAU09941 standard; protein: 402 AA.
 AC AAU09941;
 DT 09-APR-2002 (first entry)
 DE Human P-selectin glycoprotein ligand-1 (PSGL-1) protein sequence.
 XX
 XX
 XX Human; P-selectin glycoprotein ligand-1; PSGL-1, anionic target protein;
 KM protein purification; immunoglobulin domain; DNA/histone complex;
 KM DNA-protein interaction; sulphated protein.
 XX
 XX Homo sapiens.

Location/Qualifiers
 FH Key 1..41
 FT Peptide /label= Signal_peptide
 FT Domain /label= 21..310
 FT /label= Extracellular_domain
 FT Cleavage-site /label= 38..41
 FT /label= PACR_cleavage_site
 FT Protein /label= 42..402
 FT /label= Mature_PSGR-1
 FT Region 55..267
 FT /note= "Includes subdomain of 15 decameric repeats of
 ten amino acid consensus sequence"

FT Modified-site 65..67 /note="Asn is N-glycosylated"
 FT Modified-site 111..113 /note="Asn is N-glycosylated"
 FT Modified-site 292..294 /note="Asn is N-glycosylated"
 FT Domain 311..332 /label="Asn is N-glycosylated"
 FT Domain 333..402 /label="Transmembrane_domain"
 FT Domain /label="Intracellular_cytoplasmic_domain"
 PN MO200172769-A2.
 XX
 XX 04-OCT-2001.
 XX
 XX 27-MAR-2001; 2001MO-US09815.
 XX
 XX 27-MAR-2000; 2000US-193351P.
 XX
 XX (GEMV) GENNETICS INST INC.
 PA
 PI Coffman JL, Foster WB, Germann BJ, Sun S, Robinson JJ;
 XX
 XX WPI; 2001-616464/71.
 DR
 XX
 XX Purifying highly anionic target proteins from impurities and
 PT DNA/histone complexes in sample, by contacting sample to substrate that
 PT binds target, while impurities are washed off and the complexes are
 PT dissociated -
 PS Disclosure; Page 30-31; 31pp; English.
 XX
 XX The present invention provides a method for purifying highly anionic
 CC target proteins from proteaceous and non-proteaceous impurities
 CC and DNA/histone complexes in the sample. The method of the invention
 CC involves contacting the sample with a substrate that reversibly binds
 CC charged molecules, and purifying, under conditions such that the
 CC impurities are not bound to or washed off the substrate. The sample is
 CC then dissociated, while the target protein remains bound. The method is
 CC useful for isolating and purifying highly anionic target proteins e.g.
 CC sulphated proteins, from proteaceous and non-proteaceous impurities,
 CC and DNA/histone complexes in the sample. The present sequence is of
 CC the human P-selectin glycoprotein ligand-1 (PSGL-1) protein used in the
 CC invention. The ten amino acid consensus repeat of the decameric
 CC subdomain of this protein is indicated in AAU09942.
 CC
 CC
 XX
 XX Sequence 402 AA;
 SQ
 Query Match 100.0%; Score 2030; DB 22; Length 402;
 Best Local Similarity 100.0%; Pred. No. 2,3e-149;
 Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 301 APDHSIVKQCLLAIIILAVLVAITFFVCTVAVLAVLSRKGHMYPVRYNSPTMWCISISLLP 360
 QY 361 DGGEGPSATFANGGLSKAKSPGLTPPREDRREGDDTLTNSFLP 402
 DB 361 DGGEGPSATFANGGLSKAKSPGLTPPREDRREGDDTLTNSFLP 402
 RESULT 9
 AAU79653
 ID AAU79653 standard; Protein; 402 AA.
 AC AAU79653;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Human P-selectin glycoprotein ligand-1 (PSGL-1).
 XX
 KW Human; stenosis; restenosis; vascular injury; cardiovascular disease;
 KW soluble P-selectin glycoprotein ligand-1; PSGL-1; vascular remodeling;
 KW neointimal formation; coronary artery disease; peripheral artery;
 KW carotid artery; percutaneous transluminal coronary angioplasty; PTCA;
 KW stent implantation; leukocyte recruitment; cell to cell adhesion;
 KW blood vessel; vasotropic; cardiatic.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..41
 FT Protein /label= Signal_peptide
 FT /label= 402
 FT /label= Mature_PSGL-1
 FT Region 42..316
 FT /note="This region is specifically claimed in claim 24"
 FT
 XX
 XX MO200222820-A1.
 XX
 XX 21-MAR-2002.
 XX
 XX 12-SEP-2000; 2000MO-US25007.
 XX
 XX 12-SEP-2000; 2000MO-US25007.
 XX
 XX (GEMV) GENNETICS INST INC.
 PA (MONT-) MONTREAL HEART INST.
 PA
 XX Kumar A, Schnab RG, Tanguay J, Methi Y;
 PI WPI; 2002-351889/38.
 XX
 DR N-PSDB; ABR48715.
 DR
 XX
 PT Modulating stenosis or restenosis in subject having vascular injury or
 PT cardiovascular disease, comprises administering P-selectin antagonist
 PT to the subject -
 PS Claim 24; Page 57-59; 66pp; English.
 XX
 XX The present invention relates to methods and compositions for modulating
 CC stenosis or restenosis in a subject having vascular injury or
 CC cardiovascular disease. The method comprises administering a P-selectin
 CC antagonist such as soluble P-selectin glycoprotein ligand-1 (PSGL-1) to
 CC the subject. The method is useful for preventing, inhibiting or treating
 CC stenosis or restenosis in a subject, preferably human, having vascular
 CC injury or coronary artery disease. Restenosis is characterized by
 CC constrictive vascular remodeling or neointimal formation. The vascular
 CC injury or cardiovascular disease affects a coronary artery, or peripheral
 CC artery, preferably a carotid artery. The vascular injury results from
 CC angioplasty (percutaneous transluminal coronary angioplasty (PTCA)), or
 CC implantation of a stent or stents. The method is also useful for
 CC modulating leukocyte recruitment, inhibiting cell to cell adhesion, or
 CC inhibiting cell adhesion to blood vessels in a subject. The present
 CC sequence represents human PSGL-1.

SQ Sequence 402 AA;
 Query Match 100.0%; Score 2030; DB 23; Length 402;
 Best Local Similarity 100.0%; Pred. No. 2,3e-149;
 Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPQLILLIILGPGNSLQIMDTWADAEKALGPILARDRQATEYEVLDYDFLPETEP 60
 |||
 DB 1 MPQLILLIILGPGNSLQIMDTWADAEKALGPILARDRQATEYEVLDYDFLPETEP 60
 |||

QY 61 EMLRNSDTPPLTGPPTPESTVVEPARRSTGLDAGAVTELTTLANMGNLSTDSAME 120
 |||
 DB 61 EMLRNSDTPPLTGPPTPESTVVEPARRSTGLDAGAVTELTTLANMGNLSTDSAME 120
 |||

QY 121 IOTTOPAAATEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPGTE 180
 |||
 DB 121 IOTTOPAAATEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPGTE 180
 |||

QY 181 AOTTAPAAEAQTTPAAMEAQTTPPAAMEAQTTPAAMEAQTTPAAMEAQTTPAATEAQTTPATE 240
 |||
 DB 181 AOTTAPAAEAQTTPAAMEAQTTPPAAMEAQTTPAAMEAQTTPAAMEAQTTPAATEAQTTPATE 240
 |||

QY 241 AOTTPLAAMEALSTEPSATEALSMEPTTKRGLFIPSSVSVTHKGI PMAASNL SVNYPVG 300
 |||
 DB 241 AOTTPLAAMEALSTEPSATEALSMEPTTKRGLFIPSSVSVTHKGI PMAASNL SVNYPVG 300
 |||

QY 301 APDHISVKOCLLAILLILAVATTFVCTVVLAVRSLRKGHWYVARNYSPTENWCISLILP 360
 |||
 DB 301 APDHISVKOCLLAILLILAVATTFVCTVVLAVRSLRKGHWYVARNYSPTENWCISLILP 360
 |||

QY 361 DGGEGSATANGLSKAKSPGLTPPEPREDEGDDLTLSFLP 402
 |||
 DB 361 DGGEGSATANGLSKAKSPGLTPPEPREDEGDDLTLSFLP 402
 |||

RESULT 10
 AAG77945 ID AAG77945 standard; Protein; 402 AA.
 AC AAG77945;
 XX
 DT 31-JAN-2002 (first entry)
 DE Human P-selectin glycoprotein ligand-1.
 XX
 KM Human; PSEL-1; P-selectin glycoprotein ligand-1; thrombosis;
 KM thrombolytic; anticoagulant; cardiovascular disease; atherosclerosis;
 KM hypertension; angioplasty.
 XX
 OS Homo sapiens.
 XX
 PN MO200175107-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 02-APR-2001; 2001MO-US10622.
 PR 31-MAR-2000; 2000US-193787P.
 XX
 PA (GENY) GENETICS INST INC.
 XX
 PI Eppihimer MJ, Schaub RG, Harris AS;
 XX
 DR WPI: 2002-010790/01.
 DR N-PSDB; AAH77236.
 XX
 PT Modulating (e.g. preventing, inhibiting or treating) thrombosis,
 PT comprises administering a P-selectin antagonist to subject, e.g. human
 PT with cardiovascular disorder (e.g. atherosclerosis or hypertension) -
 XX
 PS Claim 13; Page 69-70; 73pp; English.
 CC The sequence represents human P-selectin glycoprotein ligand-1 (PSEL-1).

CC The invention relates to a novel method for treating or inhibiting
 CC thrombosis in a subject comprising administering a composition comprising
 CC an effective amount of a P-selectin antagonist. The polypeptides and
 CC polynucleotides of the invention have thrombolytic, and anticoagulant
 CC activity. The method is useful for treating or inhibiting thrombosis in a
 CC subject, particularly a human. The method is also useful for inhibiting
 CC cell adhesion to blood vessels in a subject, increasing the movement of
 CC cells relative to blood vessels in a subject, or inhibiting the effect of
 CC a thrombus-inducing agent in a subject. The subject is at risk for
 CC thrombosis, e.g. a subject suffering from a cardiovascular disease or
 CC disorder (e.g. atherosclerosis or hypertension), or a subject who has
 CC undergone cardiovascular or general vascular procedures or intervention
 CC such as angioplasty of any vessel.
 XX

SQ Sequence 402 AA;
 Query Match 100.0%; Score 2030; DB 23; Length 402;
 Best Local Similarity 100.0%; Pred. No. 2,3e-149;
 Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPQLILLIILGPGNSLQIMDTWADAEKALGPILARDRQATEYEVLDYDFLPETEP 60
 |||
 DB 1 MPQLILLIILGPGNSLQIMDTWADAEKALGPILARDRQATEYEVLDYDFLPETEP 60
 |||

QY 61 EMLRNSDTPPLTGPPTPESTVVEPARRSTGLDAGAVTELTTLANMGNLSTDSAME 120
 |||
 DB 61 EMLRNSDTPPLTGPPTPESTVVEPARRSTGLDAGAVTELTTLANMGNLSTDSAME 120
 |||

QY 121 IOTTOPAAATEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPGTE 180
 |||
 DB 121 IOTTOPAAATEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPGTE 180
 |||

QY 181 AOTTAPAAEAQTTPAAMEAQTTPPAAMEAQTTPAAMEAQTTPAAMEAQTTPAATEAQTTPATE 240
 |||
 DB 181 AOTTAPAAEAQTTPAAMEAQTTPPAAMEAQTTPAAMEAQTTPAAMEAQTTPAATEAQTTPATE 240
 |||

QY 241 AOTTPLAAMEALSTEPSATEALSMEPTTKRGLFIPSSVSVTHKGI PMAASNL SVNYPVG 300
 |||
 DB 241 AOTTPLAAMEALSTEPSATEALSMEPTTKRGLFIPSSVSVTHKGI PMAASNL SVNYPVG 300
 |||

QY 301 APDHISVKOCLLAILLILAVATTFVCTVVLAVRSLRKGHWYVARNYSPTENWCISLILP 360
 |||
 DB 301 APDHISVKOCLLAILLILAVATTFVCTVVLAVRSLRKGHWYVARNYSPTENWCISLILP 360
 |||

QY 361 DGGEGSATANGLSKAKSPGLTPPEPREDEGDDLTLSFLP 402
 |||
 DB 361 DGGEGSATANGLSKAKSPGLTPPEPREDEGDDLTLSFLP 402
 |||

RESULT 11
 AAR53965 ID AAR53965 standard; Protein; 402 AA.
 AC AAR53965;
 XX
 DT 14-OCT-1994 (first entry)
 DE P-selectin ligand.
 XX
 KM P-selectin ligand; glycoprotein; inflammation;
 KM intercellular adhesion; promyelocyte; Hle0; pM71; PL85;
 KM alpha-1,3/alpha-1,4-fucosyltransferase; PACSOL;
 KM paired basic amino acid converting enzyme.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH Peptide 1..41
 FT /label= Sig_peptide
 FT Domain 21..310
 FT /label= extracellular domain
 FT Misc-difference 46
 FT /note= "potential sulfation site"

FT Misc-difference 48 /note= "potential sulfation site"
 FT Misc-difference 51 /note= "potential sulfation site"
 FT Misc-difference 65 /note= "potential N-linked glycosylation site"
 FT Misc-difference 111 /note= "potential N-linked glycosylation site"
 FT Misc-difference 292 /note= "potential N-linked glycosylation site"
 FT Domain 311..332 /label= "transmembrane domain"
 FT Domain 333..402 /label= "cytoplasmic domain"
 XX MO9410309-A.
 XX 11-MAY-1994.
 XX 22-OCT-1993; 93WO-US10168.
 XX 23-OCT-1992; 92US-0965662.
 XX 26-AUG-1993; 93US-0112608.
 XX (GENW) GENETICS INST INC.
 XX Chang X, Larsen GR, Sako DS, Veldman GM;
 PI MPI: 1994-167466/20.
 DR N-PSDB; Q-63953.
 XX
 XX DNA encoding P-selectin ligand glycoprotein - for use in
 PT treating inflammatory disease characterized by P-selectin
 PR mediated intercellular adhesion
 XX
 XX Disclosure; Page 48-50; 64pp; English.
 XX
 XX CDNA of sequence AA063953, encoding P-selectin ligand, was derived
 CC from promyelocyte line HL60 clone PMT21:PL85. The ligand comprises
 CC residues 1-402, 1-310 (sol.), 42-402 (mature) or 42-310 (sol.).
 CC mature) of sequence AAR53965. An allelic variant (AA063954, AAR53966)
 CC was derived from human placenta. Co-transfection of a mammalian host
 CC cell with DNA encoding residues 42-310, plus DNA encoding a soluble
 CC alpha-1,3/alpha-1,4-fucosyltransferase and DNA encoding a soluble
 CC form of a paired basic amino acid converting enzyme (PACBSOL).
 CC encoded by sequence AA063957, allowed production of sol. mature
 CC P-selectin ligand glycoprotein.
 CC
 CC
 CC Sequence 402 AA;
 SQ

QY 301 APDHSVKQCLLAIIILAVATIFPVCTVVAVRLSRKGMYPVRYSPTEWVCISLSP 360
 DB 301 APDHSVKQCLLAIIILAVATIFPVCTVVAVRLSRKGMYPVRYSPTEWVCISLSP 360
 QY 361 DGGEGPSATANGSLSKAKSPGLTPPPRRRREDDDLTHSFLP 402
 DB 361 DGGEGPSATANGSLSKAKSPGLTPPPRRRREDDDLTHSFLP 402

RESULT 12
 AAW72696
 ID AAW72696 standard; Protein; 412 AA.
 XX AAW72696;
 XX
 DT 08-JAN-1999 (first entry)
 XX
 XX Human glycoprotein P-selectin ligand protein.
 DE Human glycoprotein P-selectin ligand protein.
 XX Human; glycoprotein; P-selectin ligand protein; inflammatory disease;
 KW intracellular domain; E-selectin-mediated intercellular adhesion.
 XX Homo sapiens.
 OS
 XX US5827817-A.
 PN
 XX 27-OCT-1998.
 PD
 XX 07-JUN-1995; 95US-0477254.
 PF
 XX 25-APR-1995; 95US-0428734.
 PR 23-OCT-1992; 92US-0965662.
 PR 26-AUG-1993; 93US-0112608.
 PR 22-OCT-1993; 93WO-US10168.
 PR 28-APR-1994; 94US-0235398.
 PR 30-SEP-1994; 94US-0316305.
 PR 07-JUN-1995; 95US-0477254.
 XX
 PA (GENW) GENETICS INST INC.
 XX
 XX Camphausen R, Chang X, Cumming D, Kumar R, Larsen GR;
 PI Sako DS, Shaw G, Veldman GM;
 PT MPI: 1998-594030/50.
 DR N-PSDB; AAV67131.
 XX
 XX Soluble P-selectin ligand glycoprotein - with truncated protein
 PT sequence lacking intracellular domain
 XX
 XX Disclosure; Column 49-52; 66pp; English.
 PS
 XX The present sequence represents a human glycoprotein having P-selectin
 CC ligand (PSL) activity. The present invention describes a glycoprotein
 CC having PSL activity which comprises a carbohydrate and a protein. The
 CC protein includes amino acids 42-60 of AAW72695, but lacks the
 CC intracellular domain and optionally the transmembrane domain of mature
 CC PSL protein, and the carbohydrate contains sialyl Lewis X. The isolated
 CC PSL protein may be useful in treating conditions characterized by P- or
 CC E-selectin mediated intercellular adhesion. Such conditions include
 CC myocardial infarction, bacterial or viral infection, metastatic
 CC conditions, inflammatory disorders, asthma, emphysema, thermal injury,
 CC multiple sclerosis, diabetes, Reynaud's syndrome, neurophilic
 CC dermatosis, Grave's disease, glomerulonephritis, gingivitis,
 CC periodontitis, haemolytic uremic syndrome, ulcerative colitis, Crohn's
 CC disease, necrotizing enterocolitis, and cytokine-induced toxicity. The
 CC isolated PSL protein may also be useful in organ transplantation, both to
 CC prepare organs for transplantation and to quell organ transplant
 CC rejection. The isolated PSL may be used to treat hemodialysis and
 CC leukopheresis patients. Additionally, isolated PSL protein may be used as
 CC an antineoplastic agent. The isolated PSL protein may be used itself as
 CC an inhibitor of P- or E-selectin-mediated intercellular adhesion or to
 CC design inhibitors of P- or E-selectin-mediated intercellular adhesion.

XX Sequence 412 AA;

Query Match 99.3%; Score 2015; DB 19; Length 412; Best Local Similarity 97.6%; Pred. No. 3.4e-148; Matches 402; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 MPEQLLILLLILGPGNSLQIMDTWADAEKALGPLLARDRROATEYEVLDYDFLPETEP 60
1 MPEQLLILLLILGPGNSLQIMDTWADAEKALGPLLARDRROATEYEVLDYDFLPETEP 60
DB 1 MPEQLLILLLILGPGNSLQIMDTWADAEKALGPLLARDRROATEYEVLDYDFLPETEP 60
QY 61 EMLRNSTDTTPTLPGPTPESTVVEPARRSTGLDAGAVTELTTELANNGNLSTDSAA 120
61 EMLRNSTDTTPTLPGPTPESTVVEPARRSTGLDAGAVTELTTELANNGNLSTDSAA 120
DB 61 EMLRNSTDTTPTLPGPTPESTVVEPARRSTGLDAGAVTELTTELANNGNLSTDSAA 120
QY 121 IOTTQPA-----TEAOTPLAATAEOTRLTATEAOTPLAATAEOTTPPATE 170
121 IOTTQPA-----TEAOTPLAATAEOTRLTATEAOTPLAATAEOTTPPATE 170
DB 121 IOTTQPA-----TEAOTPLAATAEOTRLTATEAOTPLAATAEOTTPPATE 180
QY 171 AOTTQPTGLEAOTTAPAMEAOTTAPAMEAOTTAPAMEAOTTAPAMEAOTTAPATE 230
171 AOTTQPTGLEAOTTAPAMEAOTTAPAMEAOTTAPAMEAOTTAPAMEAOTTAPATE 230
DB 181 AOTTQPTGLEAOTTAPAMEAOTTAPAMEAOTTAPAMEAOTTAPAMEAOTTAPATE 240
QY 231 AOTTQPTATAEOTTPLAAMEALSTEPSATEALSMEPTTKRGLPIPSVSSVTHKGI 290
231 AOTTQPTATAEOTTPLAAMEALSTEPSATEALSMEPTTKRGLPIPSVSSVTHKGI 290
DB 241 AOTTQPTATAEOTTPLAAMEALSTEPSATEALSMEPTTKRGLPIPSVSSVTHKGI 300
QY 291 SNLSVNVYPGAPPHISVKOCLLAILLIALVATTFVCTVVAVRLSRKGMYPVRNYSPT 350
291 SNLSVNVYPGAPPHISVKOCLLAILLIALVATTFVCTVVAVRLSRKGMYPVRNYSPT 350
DB 301 SNLSVNVYPGAPPHISVKOCLLAILLIALVATTFVCTVVAVRLSRKGMYPVRNYSPT 360
QY 351 EWCYSSLLPDDGEGPSATANGLSKAKSPGLTPBEREDREGDDLTHSFLP 402
351 EWCYSSLLPDDGEGPSATANGLSKAKSPGLTPBEREDREGDDLTHSFLP 402
DB 361 EWCYSSLLPDDGEGPSATANGLSKAKSPGLTPBEREDREGDDLTHSFLP 412

RESULT 13
AAW56346
ID AAW56346 standard; Protein; 412 AA.

AC AAW56346;
DT 27-JUL-1998 (first entry)
DE P-selectin ligand protein.
KW P-selectin ligand protein; glycoprotein; human; calcium-dependent;
intercellular adhesion; E-selectin; L-selectin; myocardial infarction;
inflammation; Crohn's disease; transplant rejection; rejection.
OS Homo sapiens.
PN K09808949-A1.
PD 05-MAR-1998.
PF 29-AUG-1997; 97WC-US14159.
PR 30-AUG-1996; 96US-0713556.
PA (GEMV) GENETICS INST INC.
PI Chang X, Gunning D, Kumar R, Larsen GR, Sako DS;
Shaw G, Veldman GM;
XX WPI; 1998-179434/16.
DR N-PSDB; AAV22850.
RT New isolated DNA encoding fusion protein including P-selectin ligand
fragment - to direct second component, e.g. cytokine to sites of
PT selectin expression, used e.g. to stimulate bone and cartilage
formation
XX

PS Disclosure; Pages 68-69; 128pp; English.

CC The present sequence represents a P-selectin ligand protein. The
CC P-selectin ligand is a glycoprotein. The protein binds in a
CC calcium-dependent manner to P-selectin which is present on the surface
CC of cells. The P-selectin ligand can be used to treat a wide variety of
CC conditions characterized by intercellular adhesion involving P-, E- or
CC L-selectins, e.g. myocardial infarction, infection, metastases,
CC inflammation, Crohn's disease, and to prevent transplant rejection. It
CC can also be used to raise specific antibody (useful therapeutically as
CC inhibitors of adhesion or for immunodiagnosis of inflammation and cancer)
CC or to screen for selective inhibitors. Fusion proteins of the P-selectin
CC ligand protein are used to treat conditions related to selectin
CC expression, e.g. fusion with interleukin-11 to treat bone marrow
CC endothelial cells to stimulate megakaryocyte progenitors, bone
CC morphogenic protein (BMP) to stimulate bone or cartilage formation,
CC or other cytokines to sites of inflammation.
XX

XX Sequence 412 AA;
Query Match 99.3%; Score 2015; DB 19; Length 412;
Best Local Similarity 97.6%; Pred. No. 3.4e-148;
Matches 402; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 MPEQLLILLLILGPGNSLQIMDTWADAEKALGPLLARDRROATEYEVLDYDFLPETEP 60
1 MPEQLLILLLILGPGNSLQIMDTWADAEKALGPLLARDRROATEYEVLDYDFLPETEP 60
DB 1 MPEQLLILLLILGPGNSLQIMDTWADAEKALGPLLARDRROATEYEVLDYDFLPETEP 60
QY 61 EMLRNSTDTTPTLPGPTPESTVVEPARRSTGLDAGAVTELTTELANNGNLSTDSAA 120
61 EMLRNSTDTTPTLPGPTPESTVVEPARRSTGLDAGAVTELTTELANNGNLSTDSAA 120
DB 61 EMLRNSTDTTPTLPGPTPESTVVEPARRSTGLDAGAVTELTTELANNGNLSTDSAA 120
QY 121 IOTTQPA-----TEAOTPLAATAEOTRLTATEAOTPLAATAEOTTPPATE 170
121 IOTTQPA-----TEAOTPLAATAEOTRLTATEAOTPLAATAEOTTPPATE 170
DB 121 IOTTQPA-----TEAOTPLAATAEOTRLTATEAOTPLAATAEOTTPPATE 180
QY 171 AOTTQPTGLEAOTTAPAMEAOTTAPAMEAOTTAPAMEAOTTAPAMEAOTTAPATE 230
171 AOTTQPTGLEAOTTAPAMEAOTTAPAMEAOTTAPAMEAOTTAPAMEAOTTAPATE 230
DB 181 AOTTQPTGLEAOTTAPAMEAOTTAPAMEAOTTAPAMEAOTTAPAMEAOTTAPATE 240
QY 231 AOTTQPTATAEOTTPLAAMEALSTEPSATEALSMEPTTKRGLPIPSVSSVTHKGI 290
231 AOTTQPTATAEOTTPLAAMEALSTEPSATEALSMEPTTKRGLPIPSVSSVTHKGI 290
DB 241 AOTTQPTATAEOTTPLAAMEALSTEPSATEALSMEPTTKRGLPIPSVSSVTHKGI 300
QY 291 SNLSVNVYPGAPPHISVKOCLLAILLIALVATTFVCTVVAVRLSRKGMYPVRNYSPT 350
291 SNLSVNVYPGAPPHISVKOCLLAILLIALVATTFVCTVVAVRLSRKGMYPVRNYSPT 350
DB 301 SNLSVNVYPGAPPHISVKOCLLAILLIALVATTFVCTVVAVRLSRKGMYPVRNYSPT 360
QY 351 EWCYSSLLPDDGEGPSATANGLSKAKSPGLTPBEREDREGDDLTHSFLP 402
351 EWCYSSLLPDDGEGPSATANGLSKAKSPGLTPBEREDREGDDLTHSFLP 402
DB 361 EWCYSSLLPDDGEGPSATANGLSKAKSPGLTPBEREDREGDDLTHSFLP 412

RESULT 14
AAV29778
ID AAV29778 standard; Protein; 412 AA.

AC AAV29778;
DT 04-NOV-1999 (first entry)
DE Human P-selectin ligand protein SEQ ID NO:4.
KW Human; P-selectin ligand; glycoprotein; fusion protein; infection;
inflammation; intercellular adhesion; ulcerative colitis; asthma;
diabetes; transplant rejection; myocardial infarction; thermal injury;
metastatic condition; autoimmune thyroiditis; multiple sclerosis;
Raynaud's syndrome; neutrophilic dermatosis; Sweet's syndrome;
KW Gravel's disease; glomerulonephritis; glomerulitis; periodontitis;
KW Crohn's disease; necrotizing enterocolitis.
OS Homo sapiens.
XX

MN WO9943834-A2.
 XX 02-SEP-1999.
 XX 25-FEB-1999; 99WO-US04302.
 XX 27-FEB-1998; 98US-0032080.
 XX (GEMV) GENETICS INST INC.
 XX
 XX Camphausen R, Chang X, Cumming D, Davis M, Kumar R;
 PI Larsen GR, Sako DS, Shaw G, Veldman GM;
 DR WPI; 1999-527628/44.
 DR N-PSDB; AAZ08860.
 XX
 XX New P-selectin ligand fusion proteins, used for treating e.g.
 PT inflammation, infections, asthma, diabetes, ulcerative colitis or
 PT transplant rejection
 XX
 XX Disclosure; Page 110-111; 145pp; English.
 XX
 CC The present invention describes P-selectin ligand fusion proteins
 CC comprising amino acids 42-60, 42-402, 42-310, 42-88, 42-118 or 42-189
 CC of the P-selectin ligand protein. The fusion proteins comprising a
 CC P-selectin ligand act as ligands for P-selectin on human endothelial
 CC cells and platelets. The isolated P-selectin ligand proteins may be
 CC useful in treating conditions characterized by P-, E- or L-selectin
 CC mediated intercellular adhesion e.g. myocardial infarction, bacterial
 CC or viral infection, metabolic conditions, inflammatory disorders,
 CC thermal injury such as burns or frostbite, autoimmune thyroiditis,
 CC experimental allergic encephalomyelitis, multiple sclerosis, multiple
 CC organ injury syndrome secondary to trauma, diabetes, Reynaud's syndrome,
 CC neutrophilic dermatosis (Sweet's syndrome), inflammatory bowel disease,
 CC Grave's disease, glomerulonephritis, gingivitis, periodontitis,
 CC haemolytic uraemic syndrome, ulcerative colitis, Crohn's disease,
 CC necrotising enterocolitis, granulocyte transfusion associated syndrome,
 CC or cytokine-induced toxicity. Isolated P-selectin ligand proteins may
 CC also be useful in organ transplantation, both to prepare organs for
 CC transplantation and to quell organ transplant rejection. P-selectin
 CC ligand proteins may be used to treat haemodialysis and leukopheresis
 CC patients or used as an antimetabolic agent. The fusion proteins can
 CC also be used to treat a condition which is affected by the protein to
 CC which the P-selectin ligand protein is fused. The fusion proteins can be
 CC used for the production of antibodies for use in therapy, detection,
 CC diagnosis and drug screening. The present sequence represents the human
 CC P-selectin ligand protein, as given in the present invention.
 CC
 CC Sequence 412 AA;
 SQ
 Query Match 99.3%; Score 2015; DB 20; Length 412;
 Best Local Similarity 97.6%; Pred. No. 3.4e-148;
 Matches 402; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 291 SNLSVNYVGVADPDHISVYQCLIAIILIALVATIFVCTVAVLAVRLSRKGMYPVRYNSPT 350
 DB 301 SNLSVNYVGVADPDHISVYQCLIAIILIALVATIFVCTVAVLAVRLSRKGMYPVRYNSPT 360
 QY 351 EMWCISSLLIPDGGEGPSATANGSLSKAKSPGILTPPREBREDDDLTLSHFLP 402
 DB 361 EMWCISSLLIPDGGEGPSATANGSLSKAKSPGILTPPREBREDDDLTLSHFLP 412
 RESULT 15
 AA081967
 ID AA081967 standard; Protein; 412 AA.
 XX
 XX AA081967;
 XX
 DT 05-FEB-1999 (first entry)
 XX
 DE Human P selectin ligand glycoprotein #2.
 DE
 KW P selectin; ligand; glycoprotein; human; screening assay; inhibitor;
 KW E selectin; intercellular adhesion; inflammatory disease; treatment.
 XX
 XX Homo sapiens.
 OS
 XX
 PN US5843707-A.
 XX
 PD 01-DEC-1998.
 XX
 PF 25-APR-1995; 95US-0428734.
 XX
 XX 25-APR-1995; 95US-0428734.
 XX
 PR 23-OCT-1992; 92US-0965662.
 PR 26-AUG-1993; 93US-0112608.
 PR 22-OCT-1993; 93WO-US10168.
 PR 28-APR-1994; 94US-0235398.
 PR 30-SEP-1994; 94US-0316305.
 XX
 XX (GEMV) GENETICS INST INC.
 XX
 PI Chang X, Cumming D, Kumar R, Larsen GR, Sako DS;
 PI Shaw G, Veldman GM;
 DR WPI; 1999-044577/04.
 DR N-PSDB; AAV64998.
 XX
 PT DNA encoding P-selectin ligand protein - useful for producing
 PT recombinant protein
 XX
 PS Disclosure; Column 53-56; 67pp; English.
 XX
 CC This sequence represents a novel human P selectin ligand glycoprotein
 CC which can be used in screening assays for inhibitors of P- or
 CC E-selectin-mediated intercellular adhesion. The protein can be used
 CC in the treatment of inflammatory disease.
 CC
 CC Sequence 412 AA;
 SQ
 Query Match 99.3%; Score 2015; DB 20; Length 412;
 Best Local Similarity 97.6%; Pred. No. 3.4e-148;
 Matches 402; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

```

QY 171 AOTTQPTGLEAQTTPAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPAPEATE 230
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 AOTTQPTGLEAQTTPAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPAPEATE 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 231 AOTTQPTATEAQTTPPLAMEALSTEBGATEALSMETTKRGLPIFESVSSTHKGIPMAA 290
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 AOTTQPTATEAQTTPPLAMEALSTEBGATEALSMETTKRGLPIFESVSSTHKGIPMAA 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 291 SNLSVNYPVGAPDHISVKOCLLAILLAVATTIFVCTVVLAVRLSRKGMYPVRNYSPT 350
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 SNLSVNYPVGAPDHISVKOCLLAILLAVATTIFVCTVVLAVRLSRKGMYPVRNYSPT 360
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 351 EMWCISLPLPDGGRGPPSANGGLSKAKSPGLTPEPREDRGGDDLTLHSTFLP 402
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 EMWCISLPLPDGGRGPPSANGGLSKAKSPGLTPEPREDRGGDDLTLHSTFLP 412
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    
```

Search completed: July 14, 2003, 06:12:23
 Job time : 77 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

QM protein - protein search, using sw model

Run on: July 14, 2003, 06:13:30 ; Search time 53 Seconds
(without alignments)
683.033 Million cell updates/sec

Title: US-09-825-580-2
Perfect score: 2030
Sequence: 1 MFLQLLLILLKPGNSLQL.....TPRPREDREGDDLTLHSFUP 402

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep.*
 - 2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep.*
 - 6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep.*
 - 10: /cgn2_6/ptodata/1/pubppaa/US05_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubppaa/US10_PUB.pep.*
 - 12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
 - 13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep.*
 - 14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2030	100.0	402	9 US-10-163-853-1	Sequence 1, App11
2	2030	100.0	402	9 US-10-211-786-2	Sequence 1, App11
3	2030	100.0	402	10 US-09-819-157-1	Sequence 1, App11
4	465	22.9	313	9 US-10-211-786-4	Sequence 4, App11
5	209.5	10.3	688	10 US-09-864-761-36047	Sequence 36047, A
6	196.5	9.7	5179	9 US-10-025-380-1068	Sequence 1068, Ap
7	196.5	9.7	5179	10 US-09-922-217-1068	Sequence 1068, Ap
8	196.5	9.7	5179	10 US-09-833-263-1068	Sequence 1068, Ap
9	195	9.6	985	9 US-09-994-064-6	Sequence 6, App11
10	195	9.6	985	9 US-09-994-064-66	Sequence 66, App11
11	189.5	9.3	1031	10 US-09-815-242-10932	Sequence 10932, A
12	187	9.2	1367	10 US-09-801-368-10932	Sequence 108, App
13	183.5	9.0	331	10 US-09-864-761-36048	Sequence 36048, A
14	177.5	8.7	638	9 US-10-234-432-59	Sequence 59, App1
15	176.5	8.7	638	9 US-10-121-988-143	Sequence 143, App
16	173	8.5	288	10 US-09-216-393-341	Sequence 341, App
17	173	8.5	288	10 US-09-216-393-344	Sequence 344, App
18	173	8.5	538	9 US-10-234-432-58	Sequence 58, App1
19	173	8.5	727	9 US-10-234-432-57	Sequence 57, App1

Result	Seq ID	Score	DB	Length	App
20	US-10-163-853-1	171.5	8.4	727	9 US-10-234-432-30
21	US-10-163-853-1	171.5	8.4	5877	9 US-10-142-515-11
22	US-10-163-853-1	171.5	8.4	5935	9 US-10-243-243A-8
23	US-10-163-853-1	169.5	8.3	1387	9 US-10-156-761-13000
24	US-10-163-853-1	168	8.3	596	9 US-10-063-547-100
25	US-10-163-853-1	168	8.3	596	9 US-10-174-590-310
26	US-10-163-853-1	168	8.3	596	9 US-10-176-758-310
27	US-10-163-853-1	168	8.3	596	9 US-10-063-616-100
28	US-10-163-853-1	168	8.3	596	9 US-10-175-737-310
29	US-10-163-853-1	168	8.3	596	9 US-10-063-502-100
30	US-10-163-853-1	168	8.3	596	9 US-10-173-706-310
31	US-10-163-853-1	168	8.3	596	9 US-10-173-738-310
32	US-10-163-853-1	168	8.3	596	9 US-10-175-752-310
33	US-10-163-853-1	168	8.3	596	9 US-10-176-482-310
34	US-10-163-853-1	168	8.3	596	9 US-10-176-700-310
35	US-10-163-853-1	168	8.3	596	9 US-10-176-913-310
36	US-10-163-853-1	168	8.3	596	9 US-10-180-552-310
37	US-10-163-853-1	168	8.3	596	9 US-10-180-557-310
38	US-10-163-853-1	168	8.3	596	9 US-10-173-700-310
39	US-10-163-853-1	168	8.3	596	9 US-10-174-572-310
40	US-10-163-853-1	168	8.3	596	9 US-10-174-579-310
41	US-10-163-853-1	168	8.3	596	9 US-10-174-582-310
42	US-10-163-853-1	168	8.3	596	9 US-10-174-588-310
43	US-10-163-853-1	168	8.3	596	9 US-10-175-733-310
44	US-10-163-853-1	168	8.3	596	9 US-10-175-740-310
45	US-10-163-853-1	168	8.3	596	9 US-10-175-743-310

ALIGNMENTS

Query Match	Best Local Similarity	Score	DB	Length	App
100.0%	100.0%	2030	9	402	
Best	Local	Similarity	100.0%	Pred. No.	4e-137
Matches	402	Conservative	0	Mismatch	0
				Indels	0
				Gaps	0
1	MFLQLLLILLKPGNSLQ	MDTWADBAEKALGPI	LLARRRQATEYEDYFLPETBP	60	
1	MFLQLLLILLKPGNSLQ	MDTWADBAEKALGPI	LLARRRQATEYEDYFLPETBP	60	
61	EMLRNSTDTPTLGTGPTPE	STTVEPAARRSTGLDGA	VTELTTELIANNGNLSTDSAA	120	
61	EMLRNSTDTPTLGTGPTPE	STTVEPAARRSTGLDGA	VTELTTELIANNGNLSTDSAA	120	
61	EMLRNSTDTPTLGTGPTPE	STTVEPAARRSTGLDGA	VTELTTELIANNGNLSTDSAA	120	
121	ITOTTPATTEAQTTPAAT	TEAQTTLVATEAQTTPA	ATEAQTTPPAATEAQTTPG	180	
121	ITOTTPATTEAQTTPAAT	TEAQTTLVATEAQTTPA	ATEAQTTPPAATEAQTTPG	180	
121	ITOTTPATTEAQTTPAAT	TEAQTTLVATEAQTTPA	ATEAQTTPPAATEAQTTPG	180	
181	AQTTPAAMEAQTTPAAME	AQTTPAAMEAQTTPAAME	AQTTPAAMEAQTTPAAME	240	
181	AQTTPAAMEAQTTPAAME	AQTTPAAMEAQTTPAAME	AQTTPAAMEAQTTPAAME	240	
181	AQTTPAAMEAQTTPAAME	AQTTPAAMEAQTTPAAME	AQTTPAAMEAQTTPAAME	240	
241	AQTTPAAMEAQTTPAAME	AQTTPAAMEAQTTPAAME	AQTTPAAMEAQTTPAAME	300	
241	AQTTPAAMEAQTTPAAME	AQTTPAAMEAQTTPAAME	AQTTPAAMEAQTTPAAME	300	
241	AQTTPAAMEAQTTPAAME	AQTTPAAMEAQTTPAAME	AQTTPAAMEAQTTPAAME	300	

Db 241 AOTTPLAAMEALSTEPSATEALSMEPTTKRGLFIPFSVSSVTHKGIIPMAASNLISVNYPVG 300
 Qy 301 APDHISVKOCLLAILLIALVATTFFVCTVVLAVLRSKGMYPVRYNSPTMEMVCISISLLP 360
 Db 301 APDHISVKOCLLAILLIALVATTFFVCTVVLAVLRSKGMYPVRYNSPTMEMVCISISLLP 360
 Qy 361 DGGEGPSATANGLSKAKSPGLTPPEPREDEBDDLTLSHFLP 402
 Db 361 DGGEGPSATANGLSKAKSPGLTPPEPREDEBDDLTLSHFLP 402

RESULT 2

US-10-211-786-2
 ; Sequence 2, Application US/10211786
 ; Publication No. US20030083258A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Michael J. Eppihmer
 ; APPLICANT: Robert G. Schaub
 ; APPLICANT: Ronald Tuma
 ; TITLE OF INVENTION: MODULATION OF LEUKOCYTE-ENDOTHELIAL INTERACTIONS FOLLOWING ISCHEM
 ; FILE REFERENCE: 8702.0099-00000
 ; CURRENT APPLICATION NUMBER: US/10/211,786
 ; PRIORITY FILING DATE: 2002-08-02
 ; PRIOR APPLICATION NUMBER: 60/309,816
 ; PRIORITY FILING DATE: 2001-08-03
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 402
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-211-786-2

Query Match 100.0%; Score 2030; DB 9; Length 402;
 Best Local Similarity 100.0%; Pred. No. 4e-137;
 Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPLQLLLLLLLILGPGNSLQIMDTWADAEKALGPIIARDRROATEYEVYLDYDFLPETPEP 60
 Db 1 MPLQLLLLLLLILGPGNSLQIMDTWADAEKALGPIIARDRROATEYEVYLDYDFLPETPEP 60
 Qy 61 EMIRNSTDTTPLNGPSTVEPAARRSTGLDAGAVTELTTELANNGNLSTDSAAE 120
 Db 61 EMIRNSTDTTPLNGPSTVEPAARRSTGLDAGAVTELTTELANNGNLSTDSAAE 120
 Qy 121 IOTTPAATEAOTTPLAATEAOTTRLTATEAOTTPLAATEAOTTPAATEAOTTPGAE 180
 Db 121 IOTTPAATEAOTTPLAATEAOTTRLTATEAOTTPLAATEAOTTPAATEAOTTPGAE 180
 Qy 181 AOTTAPAAEAOTTAPAAEAOTTAPAAEAOTTAPAAEAOTTAPAAEAOTTAPAAEAOTTAPAAE 240
 Db 181 AOTTAPAAEAOTTAPAAEAOTTAPAAEAOTTAPAAEAOTTAPAAEAOTTAPAAEAOTTAPAAE 240
 Qy 241 AOTTPLAAMEALSTEPSATEALSMEPTTKRGLFIPFSVSSVTHKGIIPMAASNLISVNYPVG 300
 Db 241 AOTTPLAAMEALSTEPSATEALSMEPTTKRGLFIPFSVSSVTHKGIIPMAASNLISVNYPVG 300
 Qy 301 APDHISVKOCLLAILLIALVATTFFVCTVVLAVLRSKGMYPVRYNSPTMEMVCISISLLP 360
 Db 301 APDHISVKOCLLAILLIALVATTFFVCTVVLAVLRSKGMYPVRYNSPTMEMVCISISLLP 360
 Qy 361 DGGEGPSATANGLSKAKSPGLTPPEPREDEBDDLTLSHFLP 402
 Db 361 DGGEGPSATANGLSKAKSPGLTPPEPREDEBDDLTLSHFLP 402

RESULT 3

US-09-819-157-1
 ; Sequence 1, Application US/09819157
 ; Patent No. US20020132991A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Coffman, J.L., et al.
 ; TITLE OF INVENTION: METHODS FOR PURIFYING HIGHLY ANIONIC PROTEINS

FILE REFERENCE: GFN-002
 ; CURRENT APPLICATION NUMBER: US/09/819,157
 ; PRIORITY FILING DATE: 2001-03-27
 ; PRIOR APPLICATION NUMBER: US 60/193,351
 ; PRIORITY FILING DATE: 2000-03-27
 ; NUMBER OF SEQ ID NOS: 1
 ; PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 402
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 ; US-09-819-157-1

Query Match 100.0%; Score 2030; DB 10; Length 402;
 Best Local Similarity 100.0%; Pred. No. 4e-137;
 Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPLQLLLLLLLILGPGNSLQIMDTWADAEKALGPIIARDRROATEYEVYLDYDFLPETPEP 60
 Db 1 MPLQLLLLLLLILGPGNSLQIMDTWADAEKALGPIIARDRROATEYEVYLDYDFLPETPEP 60
 Qy 61 EMIRNSTDTTPLNGPSTVEPAARRSTGLDAGAVTELTTELANNGNLSTDSAAE 120
 Db 61 EMIRNSTDTTPLNGPSTVEPAARRSTGLDAGAVTELTTELANNGNLSTDSAAE 120
 Qy 121 IOTTPAATEAOTTPLAATEAOTTRLTATEAOTTPLAATEAOTTPAATEAOTTPGAE 180
 Db 121 IOTTPAATEAOTTPLAATEAOTTRLTATEAOTTPLAATEAOTTPAATEAOTTPGAE 180
 Qy 181 AOTTAPAAEAOTTAPAAEAOTTAPAAEAOTTAPAAEAOTTAPAAEAOTTAPAAEAOTTAPAAE 240
 Db 181 AOTTAPAAEAOTTAPAAEAOTTAPAAEAOTTAPAAEAOTTAPAAEAOTTAPAAEAOTTAPAAE 240
 Qy 241 AOTTPLAAMEALSTEPSATEALSMEPTTKRGLFIPFSVSSVTHKGIIPMAASNLISVNYPVG 300
 Db 241 AOTTPLAAMEALSTEPSATEALSMEPTTKRGLFIPFSVSSVTHKGIIPMAASNLISVNYPVG 300
 Qy 301 APDHISVKOCLLAILLIALVATTFFVCTVVLAVLRSKGMYPVRYNSPTMEMVCISISLLP 360
 Db 301 APDHISVKOCLLAILLIALVATTFFVCTVVLAVLRSKGMYPVRYNSPTMEMVCISISLLP 360
 Qy 361 DGGEGPSATANGLSKAKSPGLTPPEPREDEBDDLTLSHFLP 402
 Db 361 DGGEGPSATANGLSKAKSPGLTPPEPREDEBDDLTLSHFLP 402

RESULT 4

US-10-211-786-4
 ; Sequence 4, Application US/10211786
 ; Publication No. US20030083258A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Michael J. Eppihmer
 ; APPLICANT: Robert G. Schaub
 ; APPLICANT: Ronald Tuma
 ; TITLE OF INVENTION: MODULATION OF LEUKOCYTE-ENDOTHELIAL INTERACTIONS FOLLOWING ISCHEM
 ; FILE REFERENCE: 8702.0099-00000
 ; CURRENT APPLICATION NUMBER: US/10/211,786
 ; PRIORITY FILING DATE: 2002-08-02
 ; PRIOR APPLICATION NUMBER: 60/309,816
 ; PRIORITY FILING DATE: 2001-08-03
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 313
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-211-786-4

Query Match 22.9%; Score 465; DB 9; Length 313;
 Best Local Similarity 97.8%; Pred. No. 1e-25;
 Matches 89; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MPLQLLLLLLLILGPGNSLQIMDTWADAEKALGPIIARDRROATEYEVYLDYDFLPETPEP 60

Db 1 MPQLLLLLLLPQSGSLQMDTWABAEKALGPLLARORRQATEYEDYDPLPETTEP 60
 QY 61 EMLRNSTDTPLTGPSTPESTTVEPARARST 91
 61 EMLRNSTDTPLTGPSTPESTTVEPARARPH 91

RESULT 5
 US-09-864-761-36047
 ; Sequence 36047, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 ; FILE REFERENCE: Aecmlca-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 36047
 ; LENGTH: 688
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURES:
 ; OTHER INFORMATION: MAP TO AB023048.1
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
 ; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.7
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
 OTHER INFORMATION: EXPRESSED IN BF474, SIGNAL = 1.4
 US-09-864-761-36047

Query Match 10.3%; Score 209.5; DB 10; Length 688;
 Best Local Similarity 24.3%; Pred. No. 4,2e-07;
 Matches 66; Conservative 46; Mismatches 125; Indels 35; Gaps 5;
 QY 57 TEPPEMLRNSDTDTPLTGPSTPESTTVEPARARSTGLDAGAVTEL-----TTELAMN 109
 Db 351 SETTTVSTTGTERT-ITRSGSETTTVTAGSETTAVTTGSETTTTSTGSETTTVSTT 409
 QY 110 GNLSTDSAAAEIOTTPAATAEQTPPLAATEAQTTLPLAATEAQTTLPLAAT 159
 Db 410 GSETTTASTADLETTTVTSRSGGTTTASTAGSETTTVTVITGSKTTTASTRGEATTVST 469
 QY 160 EAQTTPPAAT-----EAQTTPGTEAQTTPAAMEAQTTPAAMEAQTTPA 207
 Db 470 SSETTTASTGSEMTTVPPTVSETTTVSTIGSEATTTSSAAGSEATTTSTGSETTTASTA 529
 QY 208 AMEAQTTPAAMEAQTTPAATEAQTTPAATEAQTTPAAMEA-----LSTEPSATEAL 262
 Db 530 GSETTTASTAGSETTTASTAGSETTTACTGSETTSTPSAGSETTNTAFITGSSSTTASTA 589
 QY 263 SMEPTTKRGLRIFPSVSVTHKGIIPMAASNLS 294
 Db 590 SLEPTATSLTGETTTVSTTASGATRAASTTVS 621

RESULT 6
 US-10-025-380-1068
 ; Sequence 1068, Application US/10025380
 ; Publication No. US20020182191A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jianshun
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Secrist, Heather
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Meagher, Madeleine Joy
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Wang, Tongrong
 ; APPLICANT: Wang, Yugu
 ; APPLICANT: Smith, Carole L.
 ; APPLICANT: King, Gordon E.
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Clapper, Jonathan D.
 ; APPLICANT: Skeiky, Yasir A. W.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Vedvick Thomas S.
 ; APPLICANT: Carter, Darrick
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
 ; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
 ; FILE REFERENCE: 210121.471C14
 ; CURRENT APPLICATION NUMBER: US/10/025,380
 ; CURRENT FILING DATE: 2001-12-19
 ; NUMBER OF SEQ ID NOS: 1129
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1068
 ; LENGTH: 5179
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-025-380-1068
 Query Match 9.7%; Score 196.5; DB 9; Length 5179;
 Best Local Similarity 24.5%; Pred. No. 3.9e-05;
 Matches 89; Conservative 24; Mismatches 169; Indels 81; Gaps 13;
 QY 55 PTEPEPEMLRNSDTDTPLTGPSTPESTTVEPARARSTGLDAGAVTELTTTELAMN 114
 Db 3966 PTTTPTTTTPTTTTPTTPTPTPTPTPTTTTPTTTTPTTTTPTTPTPTPTTTTPTTT 4024
 QY 115 DSAAAEIOTTPAATEA-QTTPAATEAQTTLPLAATEAQTTLPLAATEAQTTPAATE 170

Db 4025 -----PTEPPTGTQTPPTTTT-----PAAMEAQTTPAAMEAQTTP-PAAMEAQTTPAAMEAQTTP 4077
 QY 171 AQTTPGTLEAQT-----APAMEAQTTPAAMEAQTTP-PAAMEAQTTPAAMEAQTTP 224
 Db 4078 TQTPPTTTPTTTTTPPTPTPTPTPTPTPTPTTTTTPPTTTPPTPTPTPTPTPTPTPTPTPTPTPTPTPT 4137
 QY 225 -----APEATEAQTTPAATEAQTTPAAMEALSTEPSATEALSMETTKRGLFIPSPVS 279
 Db 4138 VTPPTPTGTQTPPTTTTTP-----TPPTGTQTPPTTTTTPPTPTPTPTPTPTPTPTPTPTPTPT 4190
 QY 280 SVTHKGIIPMAAS-----NLSVNVYVPGAPDHSVKQCLLAILLIALVATIFVCTVLAVR 334
 Db 4191 TGTQTPPTHTSTAPPAELTTSNP-----PPESSTPQ-----ATANGLSKAKSPGLT 4222
 QY 335 LSRKGMHYEVRYNSP-TEWVCISLILPDGEGPS-----ATANGLSKAKSPGLT 383
 Db 4223 TGRS-----TSSPLTBSSTLLSTLPPAILEMTSTAPBSTPPTPAPTSTSGHHTLSPSPSTT 4275
 QY 384 PEP 386
 Db 4276 TSP 4278

RESULT 7

US-09-922-217-1068
 ; Sequence 1068, Application US/09922217
 ; Patent No. US20020076414A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Secrist, Heather
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Meagher, Madeleine Joy
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Wang, Tongcong
 ; APPLICANT: Jiang, Yugu
 ; APPLICANT: Smith, Carole Lynn
 ; APPLICANT: King, Gordon E.
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Clapper, Jonathan D.
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
 ; FILE REFERENCE: 210121.471C13
 ; CURRENT APPLICATION NUMBER: US/09/922,217
 ; CURRENT FILING DATE: 2001-08-03
 ; NUMBER OF SEQ. ID NOS: 1124
 ; SOFTWARE: FaetsEQ for Windows Version 4.0
 ; SEQ ID NO 1068
 ; LENGTH: 5179
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-922-217-1068

Query Match 9.7%; Score 196.5; DB 10; Length 5179;
 Best Local Similarity 24.5%; Pred. No. 3.9e-05;
 Matches 89; Conservative 24; Mismatches 169; Indels 81; Gaps 13;

QY 55 PETPEPMRNSTDTTPLTGPPTPESTVYEPARASTGLDAGAVTLLTTELAMGNLST 114
 Db 3966 PTTTPTTTTTVTPPTPTPTPTPTPTPTPTTTTTPPTTTPPTPTPTPTPTPTPTPTPTPTPTPTPT 4024
 QY 115 DSAMEIQTTPAATEA-OTTPAATEAQTTPAATEAQTTPAATEAQTTPAATEAQTTPAATEAQTTPAATEA 170
 Db 4025 -----PPTPTGTQTPPTTTTPTTTTTPPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 4077
 QY 171 AQTTPGTLEAQT-----APAMEAQTTPAAMEAQTTP-PAAMEAQTTPAAMEAQTTP 224
 Db 4078 TQTPPTTTPTTTTTPPTPTPTPTPTPTPTPTTTTTPPTTTPPTPTPTPTPTPTPTPTPTPTPTPTPTPT 4137
 QY 225 -----APEATEAQTTPAATEAQTTPAAMEALSTEPSATEALSMETTKRGLFIPSPVS 279
 Db 4138 VTPPTPTGTQTPPTTTTTP-----TPPTGTQTPPTTTTTPPTPTPTPTPTPTPTPTPTPTPTPT 4190

QY 280 SVTHKGIIPMAAS-----NLSVNVYVPGAPDHSVKQCLLAILLIALVATIFVCTVLAVR 334
 Db 4191 TGTQTPPTHTSTAPPAELTTSNP-----PPESSTPQ-----ATANGLSKAKSPGLT 4222
 QY 335 LSRKGMHYEVRYNSP-TEWVCISLILPDGEGPS-----ATANGLSKAKSPGLT 383
 Db 4223 TGRS-----TSSPLTBSSTLLSTLPPAILEMTSTAPBSTPPTPAPTSTSGHHTLSPSPSTT 4275
 QY 384 PEP 386
 Db 4276 TSP 4278

RESULT 8

US-09-833-263-1068
 ; Sequence 1068, Application US/09933263
 ; Patent No. US2002010547A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Clapper, Jonathan D.
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Meagher, Madeleine J.
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
 ; FILE REFERENCE: 210121.471C12
 ; CURRENT APPLICATION NUMBER: US/09/833,263
 ; CURRENT FILING DATE: 2001-04-10
 ; NUMBER OF SEQ. ID NOS: 1093
 ; SOFTWARE: FaetsEQ for Windows Version 3.0
 ; SEQ ID NO 1068
 ; LENGTH: 5179
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-833-263-1068

Query Match 9.7%; Score 196.5; DB 10; Length 5179;
 Best Local Similarity 24.5%; Pred. No. 3.9e-05;
 Matches 89; Conservative 24; Mismatches 169; Indels 81; Gaps 13;

QY 55 PETPEPMRNSTDTTPLTGPPTPESTVYEPARASTGLDAGAVTLLTTELAMGNLST 114
 Db 3966 PTTTPTTTTTVTPPTPTPTPTPTPTPTPTTTTTPPTTTPPTPTPTPTPTPTPTPTPTPTPTPTPT 4024
 QY 115 DSAMEIQTTPAATEA-OTTPAATEAQTTPAATEAQTTPAATEAQTTPAATEAQTTPAATEAQTTPAATEA 170
 Db 4025 -----PPTPTGTQTPPTTTTPTTTTTPPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 4077
 QY 171 AQTTPGTLEAQT-----APAMEAQTTPAAMEAQTTP-PAAMEAQTTPAAMEAQTTP 224
 Db 4078 TQTPPTTTPTTTTTPPTPTPTPTPTPTPTPTTTTTPPTTTPPTPTPTPTPTPTPTPTPTPTPTPTPTPT 4137
 QY 225 -----APEATEAQTTPAATEAQTTPAAMEALSTEPSATEALSMETTKRGLFIPSPVS 279
 Db 4138 VTPPTPTGTQTPPTTTTTP-----TPPTGTQTPPTTTTTPPTPTPTPTPTPTPTPTPTPTPTPT 4190
 QY 280 SVTHKGIIPMAAS-----NLSVNVYVPGAPDHSVKQCLLAILLIALVATIFVCTVLAVR 334
 Db 4191 TGTQTPPTHTSTAPPAELTTSNP-----PPESSTPQ-----ATANGLSKAKSPGLT 4222
 QY 335 LSRKGMHYEVRYNSP-TEWVCISLILPDGEGPS-----ATANGLSKAKSPGLT 383
 Db 4223 TGRS-----TSSPLTBSSTLLSTLPPAILEMTSTAPBSTPPTPAPTSTSGHHTLSPSPSTT 4275
 QY 384 PEP 386
 Db 4276 TSP 4278

RESULT 9

US-09-994-064-6
 ; Sequence 6, Application US/09994064
 ; Publication No. US20030082788A1

GENERAL INFORMATION:

APPLICANT: Wild, Martha A.
 APPLICANT: Cochran, Mark D.
 TITLE OF INVENTION: RECOMBINANT INFECTIOUS IARVINGOTRACHEITIS VIRUS
 TITLE OF INVENTION: AND USES THEREOF
 NUMBER OF SEQUENCES: 72
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Cooper & Dunham LLP
 STREET: 1185 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/994,064
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/468,190
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 39116-A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 278-0400
 TELEFAX: (212) 391-0525
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 985 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-994-064-6

Query Match 9.6%; Score 195; DB 9; Length 985;

Best Local Similarity 24.2%; Pred. No. 6.9e-06; Indels 106; Gaps 16;

Matches 93; Conservative 48; Mismatches 138; Indels 106; Gaps 16;

14 PGNLSIQMDTWAD-----AKKALGPLLAR-----DRQQTAEYEVLYDYDF 53
 338 PGATLPPFPPTAAPDPTGSPPTTYVPEPAITLILRSTSDMGFSTARATGSETL---- 393
 54 LPETPEPMKRNSTDTTPTLTPGPTP--ESTTYEPAARSTGLDAGAVTELTTELANNGN 111
 394 ---SVVQETDRLSTLSTPLPLTPGSESENTLFPPTAAGISTETPSAAHETTQOTSAETV 450
 112 LSTDSAAAEIQT-----TOPAAEAOQT--TPLAATEAOQTRRLTAATEAOQT--T 154
 451 VFTQSPSTSESEFARQSOSQEPWYFTQTPSTEQALVTOQTQIAETFALEFLTPQPSAEOMTFQOT 510
 155 PAAATEA-----QTPPAATEAOQTPTGLEAQTTPAAM-----EAQTTAPA 197
 511 PAAETEAQAQTSTTIEITQTSRSTPEPARAPSAABEVFTQSSSTVTEVFTQTPSTVYK 570
 198 AEAQTTTPA-----AMEAQTQTAM--EAQTTAPA 234
 571 TLLSSTEBALFTRTOSAGTEAFQTSSAABPTMTQSTETHTFFTOAASVYKAKAQTQST 630
 235 QP-TATBEAQT-PLAMEALSTEPSATEALSMET--TKRGLFIPFSVSVYTHKGIPTMA 290
 631 EPEVLITQSPSTSEVPEPTRLTGAEPIITQTPSAABEVYTRSSSTMPETQAOST-----PLAS 685
 291 SNLSVNV-----YVVGABDH 304
 686 QNPFTSSGTHNTERTTYVQITTPH 710

RESULT 10

US-09-994-064-66
 Sequence 66, Application US/09994064
 Publication No. US20030082788A1
 GENERAL INFORMATION:
 APPLICANT: Wild, Martha A.
 APPLICANT: Cochran, Mark D.
 TITLE OF INVENTION: RECOMBINANT INFECTIOUS IARVINGOTRACHEITIS VIRUS
 TITLE OF INVENTION: AND USES THEREOF
 NUMBER OF SEQUENCES: 72
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Cooper & Dunham LLP
 STREET: 1185 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/994,064
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/468,190
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 39116-A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 278-0400
 TELEFAX: (212) 391-0525
 INFORMATION FOR SEQ ID NO: 66:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 985 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-994-064-66

Query Match 9.6%; Score 195; DB 9; Length 985;

Best Local Similarity 24.2%; Pred. No. 6.9e-06; Indels 106; Gaps 16;

Matches 93; Conservative 48; Mismatches 138; Indels 106; Gaps 16;

14 PGNLSIQMDTWAD-----AKKALGPLLAR-----DRQQTAEYEVLYDYDF 53
 338 PGATLPPFPPTAAPDPTGSPPTTYVPEPAITLILRSTSDMGFSTARATGSETL---- 393
 54 LPETPEPMKRNSTDTTPTLTPGPTP--ESTTYEPAARSTGLDAGAVTELTTELANNGN 111
 394 ---SVVQETDRLSTLSTPLPLTPGSESENTLFPPTAAGISTETPSAAHETTQOTSAETV 450
 112 LSTDSAAAEIQT-----TOPAAEAOQT--TPLAATEAOQTRRLTAATEAOQT--T 154
 451 VFTQSPSTSESEFARQSOSQEPWYFTQTPSTEQALVTOQTQIAETFALEFLTPQPSAEOMTFQOT 510
 155 PAAATEA-----QTPPAATEAOQTPTGLEAQTTPAAM-----EAQTTAPA 197
 511 PAAETEAQAQTSTTIEITQTSRSTPEPARAPSAABEVFTQSSSTVTEVFTQTPSTVYK 570
 198 AEAQTTTPA-----AMEAQTQTAM--EAQTTAPA 234
 571 TLLSSTEBALFTRTOSAGTEAFQTSSAABPTMTQSTETHTFFTOAASVYKAKAQTQST 630
 235 QP-TATBEAQT-PLAMEALSTEPSATEALSMET--TKRGLFIPFSVSVYTHKGIPTMA 290
 631 EPEVLITQSPSTSEVPEPTRLTGAEPIITQTPSAABEVYTRSSSTMPETQAOST-----PLAS 685
 291 SNLSVNV-----YVVGABDH 304

DB 686 QNPSSGSGTGHNTPEPRRYVQTPPH 710

RESULT 11

US-09-815-242-10932
Sequence 10932, Application US/09815242
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlson, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
LENGTH: 1031
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-815-242-10932

Query Match 9.3%; Score 189.5; DB 10; Length 1031;
Best Local Similarity 22.4%; Pred. No. 1,8e-05;
Matches 64; Conservative 66; Mismatches 131; Indels 25; Gaps 8;

DB 679 MQBALIDTGTGEGSSKQLQQLLAWSD--ELLLEPILGGFTPADAQKRINQ--LRTQIKT 724
3 LQLLLLLLIIGPNSLQWMD--TWADAEKALGPLLARDRRQATEVEYLDYDFLPETPEP 60
61 EML---RNSDTTTPPLGPGTPESTYVPPARRSTGID-----AGAVTELTTELAN 108
735 ALLLVEKSTETTSNTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 794
109 MGNLSTDSAMRIQTTQTPAATTEAQTPLAATTEAQTPLAATTEAQTPLAATTEAQTPLA 168
795 TSSSSSTPSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 864
169 TEAQTTPGLEAQTAPPAAMEAQTAPPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPA 228
855 TSSSSSTPSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 914
229 TEAQTTPATE-----AQTTPLAAMEALSTESAPBAISMETTK 269
915 TSSSSSTPSTLNENSOSKQNSVYAVES--NODPNDASNS--KPSAK 958
RESULT 12
US-09-801-368-108
Sequence 108, Application US/09801368
Patent No. US20020128250A1

GENERAL INFORMATION:
APPLICANT: Busby, Robert
APPLICANT: Calli, Brian
APPLICANT: Hecht, Peter
APPLICANT: Holtzman, Doug
APPLICANT: Madsen, Kevin
APPLICANT: Maxon, Mary
APPLICANT: Milne, Todd
APPLICANT: No. US20020128250A1man, Thea
APPLICANT: Royer, John
APPLICANT: Salama, Sofie
APPLICANT: Sherman, Amir
APPLICANT: Silva, Jeff
APPLICANT: Summers, Eric
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
FILE REFERENCE: 109272.147
CURRENT APPLICATION NUMBER: US/09/801,368
PRIOR FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 09/487,558
PRIOR FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: US 60/160,587
PRIOR FILING DATE: 1999-10-20
NUMBER OF SEQ ID NOS: 440
SOFTWARE: PatentIn version 3.0
SEQ ID NO 108
LENGTH: 1367
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-801-368-108

RESULT 13

US-09-864-761-36048
Sequence 36048, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Rank, David R.
APPLICANT: Penn, Sharon G.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aemlica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04

Query Match 9.2%; Score 187; DB 10; Length 1367;
Best Local Similarity 24.0%; Pred. No. 3.8e-05;
Matches 55; Conservative 36; Mismatches 110; Indels 28; Gaps 4;

DB 76 GTPSTVPPARRSTGIDAGAVTELTTELANMGNLSTDSAMRIQTTQTPAATTEAQTTP 135
207 GTKSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSST 244
136 LAATEAQTPLAATTEAQTPLAATTEAQTPLAATTEAQTPLAATTEAQTPLAATTEA 195
245 SSTSSTSSSTTAATPTTTSCTEKRPPTTTSCTEKRPPTTTSCTEKRPPTTTSCTEK 303
196 PAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAAMEA 253
304 KTCCKTKTTTVPPTPSSSTSSAPVPPSSSTSSAPVPPSSSTSSAPVPPSSSTSSAP 363
254 TEPSTATLSMEPTTKRGLIFPVSVTHKGIIPMAASLVNRYVCGAP 302
364 TSSSSAPVTS---STSSSAPVTSSTSSSAPVPPSSSTSSSAP 409


```

PRIORITY APPLICATION NUMBER: US 60/236,359
PRIORITY FILING DATE: 2000-09-27
PRIORITY APPLICATION NUMBER: PCT/US01/00666
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00667
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00664
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00668
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 491117
SOFTWARE: Annotmax Sequence Labeling Engine vers. 1.1
SEQ ID NO 36048
LENGTH: 331
TYPE: PRN
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AB023048.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.7
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.4
US-09-864-761-36048

```

```

Query Match      9.0%; Score 183.5; DB 10; Length 331;
Best Local Similarity 25.8%; Pred. No. 1.2e-05;
Matches 61; Conservative 51; Mismatches 99; Indels 25; Gaps 6;

QY 57 TEPPMRLKNSDTPPLTGCTGPESTTVVEPARRSTGLDAGAVT-----ELTTELAM 109
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 53 SEPTMSTWSSSTWSSSTWSSSTWSSSTWSSSTWSSSTWSSSTWSSSTWSSSTWSS 111
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 110 GNSTSTSAAMEIQT-----TOPAAEATOTPLAAAE-AQTTRRLTAETAOETPLAA 159
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 112 GSEETTTASTWSSSTWSSSTWSSSTWSSSTWSSSTWSSSTWSSSTWSSSTWSS 171
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 160 EAQTTTPEATTAOQTPTGLEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPA 219
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 172 DSEITTTCTGREGSMTAVSTVTFEFTTASTTASGSEITLITASTSDSTTASTG 231
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 220 EAQTTPPEATTAOQTPTGLEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTP 268
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 232 GSEHTKATAYTTSSESTTASTASNTGLTETVFTTIGSDTPTTASTGSEHTTAV 287
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

```

GENERAL INFORMATION:
APPLICANT: Homer, Mary J.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond L.
APPLICANT: Pearing, David H.
APPLICANT: McNeill, Patricia D.
TITLE OF INVENTION: TREATMENT OF BABESIA INFECTION
FILE REFERENCE: 210121.560
CURRENT APPLICATION NUMBER: US/10/234,432
NUMBER OF SEQ ID NOS: 108
SOFTWARE: Corixa Invention Disclosure Database
SEQ ID NO 59
LENGTH: 638
TYPE: PRN
ORGANISM: Babesia sp. WAI
US-10-234-432-59

```

```

Query Match      8.7%; Score 177.5; DB 9; Length 638;
Best Local Similarity 25.5%; Pred. No. 7.3e-05;
Matches 81; Conservative 24; Mismatches 110; Indels 103; Gaps 14;

QY 42 QATTEBYLDYDPLPETPEPMLRNSTDTPLTGCPTPESTTVVEPPA-----RRTGLDPA 95
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 42 QAT-----DVAPESSDDP-----TODAQAPSPNP-TPASVAVTPERASOGSANOOQOTGA 90
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 96 GAVTVELTTELAMGNLSRDSAAE-----IOTTOPAA---128
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 91 GEGQPVLSSTEWALTVBEETVPEPKVENVVVQEAITTPPAQVPAVENVSOPTOTVAPABA 150
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 129 -----TEAQTPLA---ATEAQTTRLTATEAQTTPAATEA-----161
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 151 PQQPAQVAPQATVAGIQQAPQVATETATAEQVVAATTEVQ-MEQAAAESPAFILEEFPQ 209
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 162 -----QTPP-----PAATEAQTTPGTEAQTTPAAMEAQTTPA-----PAAEAOQT 204
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 210 VMTQTAPEVETQAPVATVESAPQOPAVAPQAPVAVPQAVQATVAGIQQAPQVAAEAQVV 269
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 205 PPAEAOQT-----TQTTMEEAQTTPAPEATEAQTTPAAMEAQTTPA-----PAAEAOQT 254
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 270 QPVPQATQTRPVAQPVVVVVAEAVQVQPVKAAQ-AQPVKDAQDQAVASVAPQATVAGIQQ 328
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 255 -----EFSATEALMEPPTK 269
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 329 AQPVVAEAAEQVQVQPVK 346
    |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

```

RESULT 15
US-10-121-988-143
Sequence 143; Application US/10121988
Publication No. US20030068327A1
GENERAL INFORMATION:
APPLICANT: Hosken, Nancy Ann
APPLICANT: McGowan, Patrick
APPLICANT: Sleath, Paul R.
APPLICANT: Moseman, Sally P.
APPLICANT: Evans, Lawrence S.
APPLICANT: Swanson, Ryan M.
APPLICANT: McNeill, Patricia D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
FILE REFERENCE: 210121.538G1
CURRENT APPLICATION NUMBER: US/10/121,988
NUMBER OF SEQ ID NOS: 183
SOFTWARE: FaestSeq for Windows Version 4.0
SEQ ID NO 143
LENGTH: 699
TYPE: PRN
ORGANISM: HSV-2
US-10-121-988-143

```

Query Match 8.7%; Score 176.5; DB 9; Length 699;
 Best Local Similarity 22.8%; Pred. No. 9.6e-05;
 Matches 95; Conservative 49; Mismatches 167; Indels 105; Gaps 17;

```

QY 24 WA-----DEARKLGLPLARDRC-ATEYEVLDYDF-----LPEPEPEMLRNST 67
DB 300 WARGALDDGYPVAPPPRRFRRLRITDEGVDPVRAARTGRRLMALTEDTSSDPTSAP 359
QY 68 DTTPLTGPCTPESTTVEPPAARSTGLDAGAVT---ELTTELAMGNLSTDSAAMEIQTT 124
DB 360 EKTPLPVSATMMAPSVDPSAAPT---APATTPPEDEMATQAATVAATPEETAV---AS 411
QY 125 QPAAATEAQTTPLAATEAQT-----TRLVATEAQTTPLAATEAQTTPPAATEAQTTOPTGL 179
DB 412 PPAIVASVSSPLLPAALAAATPGAGHTNTSSASAAKTP-PTTPAPPTTSTHATPRPTTP 470
QY 180 EAQTT-----APAAAEQTTAPAAAEQTTTPAAAEQTTQTTAAEQTTAPAEATEAQ 232
DB 471 GPOTTPPGPATPGVGSAAAPLADSPPLTA--SPPATAPGPSANVSAVAATTAATPGTRGTA 528
QY 233 TTOPT-----ATEAQTPLAAMEAL 252
DB 529 RTEPPTDKTTPHGRADAPRGSPPAPPEPHRGGEPEFEGAGDGEPEDDDSATGLAERTPN 588
QY 253 STEPSATEALSMETTRRGLFIPFSVSSVTHKGI PMAASNLVNYYPVG-APDHSVKQCL 311
DB 589 PNKPPPARPGPIRPTLPPGILGPLAPMTPR---PPAQAP-AKDMPSGPTPOHILPFWFL 643
QY 312 LA-----LILALVVTIFVCTVVLAVRL-----SRKGMVYVRYNSPTMVCVCI 355
DB 644 TASPALDILPFIISTTHHTAAVCLVALAQLMRGRAGRRRYAHPSVRY-----VCL 694

```

Search completed: July 14, 2003, 06:22:20
 Job time : 71 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: July 14, 2003, 06:11:00 ; Search time 26 Seconds
(without alignments)
454,924 Million cell updates/sec

Title: US-09-825-580-2

Perfect score: 2030

Sequence: 1 MFLQDLLLILLGPGNSIQ.....TPRPDRRGGDDTLTHSFLP 402

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

262574

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: /cgnt2_6/prodata/1/1aa/5A_COMB.pep:**
- 2: /cgnt2_6/prodata/1/1aa/5B_COMB.pep:**
- 3: /cgnt2_6/prodata/1/1aa/5A_COMB.pep:**
- 4: /cgnt2_6/prodata/1/1aa/5B_COMB.pep:**
- 5: /cgnt2_6/prodata/1/1aa/CTUS_COMB.pep:**
- 6: /cgnt2_6/prodata/1/1aa/backfill1.pep:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2030	100.0	402	US-08-477-254A-2	Sequence 2, App11
2	2030	100.0	402	US-08-472-576B-2	Sequence 2, App11
3	2030	100.0	402	US-08-428-734B-2	Sequence 2, App11
4	2030	100.0	402	US-09-063-237-1	Sequence 1, App11
5	2030	100.0	402	US-08-713-556F-2	Sequence 2, App11
6	2015	99.3	412	US-08-477-254A-4	Sequence 4, App11
7	2015	99.3	412	US-08-472-576B-4	Sequence 4, App11
8	2015	99.3	412	US-08-428-734B-4	Sequence 4, App11
9	2015	99.3	412	US-08-713-556F-4	Sequence 4, App11
10	465	22.9	264	US-08-713-556F-42	Sequence 42, App1
11	465	22.9	313	US-08-713-556F-36	Sequence 36, App1
12	463	22.8	269	US-08-713-556F-38	Sequence 38, App1
13	278	13.7	437	US-08-713-556F-40	Sequence 40, App1
14	232	11.4	42	PCT-US96-10043-3	Sequence 3, App11
15	216	10.6	42	PCT-US96-10043-13	Sequence 13, App1
16	196.5	9.7	249	US-08-700-651-15	Sequence 15, App1
17	196.5	9.7	249	US-08-928-361B-20	Sequence 20, App1
18	195.5	9.6	1721	US-08-700-651-5	Sequence 5, App1
19	195.5	9.6	1721	US-08-928-361B-6	Sequence 6, App1
20	195	9.6	985	PCT-US96-03916-6	Sequence 6, App1
21	189.5	9.3	216	PCT-US96-03916-66	Sequence 66, App1
22	187.5	9.2	786	US-08-928-361B-27	Sequence 27, App1
23	187.5	9.2	805	US-09-103-429A-3	Sequence 3, App1
24	187.5	9.2	805	US-09-103-429A-4	Sequence 4, App1
25	182.5	9.0	216	US-08-928-361B-8	Sequence 8, App1
26	182.5	9.0	1837	US-08-928-361B-5	Sequence 5, App1
27	171.5	8.4	750	US-09-165-239A-4	Sequence 4, App1

28	167	8.2	878	4	US-09-556-706B-2	Sequence 2, App11
29	165.5	8.2	451	1	US-08-287-001A-2	Sequence 2, App11
30	165.5	8.2	451	5	PCT-US95-09941-2	Sequence 2, App11
31	158.5	7.8	1848	4	US-08-296-791-6	Sequence 6, App1
32	158.5	7.8	1848	5	PCT-US95-10661A-6	Sequence 6, App1
33	158	7.8	2476	2	US-08-276-967-2	Sequence 2, App11
34	157	7.7	1048	4	US-08-171-699-10	Sequence 10, App1
35	154.5	7.6	907	3	US-08-783-774-2	Sequence 2, App11
36	154.5	7.6	907	4	US-09-328-599A-1	Sequence 1, App1
37	154.5	7.6	907	5	PCT-US95-04611A-19	Sequence 19, App1
38	154	7.6	175	3	US-08-700-651-12	Sequence 12, App1
39	154	7.6	175	3	US-08-928-361B-17	Sequence 17, App1
40	153	7.5	941	4	US-07-757-022B-14	Sequence 14, App1
41	153	7.5	1022	4	US-07-757-022B-84	Sequence 84, App1
42	153	7.5	1038	4	US-07-757-022B-74	Sequence 74, App1
43	153	7.5	1049	4	US-07-757-022B-58	Sequence 58, App1
44	153	7.5	1140	4	US-07-757-022B-104	Sequence 104, App1
45	153	7.5	1270	4	US-07-757-022B-44	Sequence 44, App1

ALIGNMENTS

RESULT 1
US-08-477-254A-2
; Sequence 2, Application US/08477254A
; Patent No. 5827817
; GENERAL INFORMATION:
; APPLICANT: Larsen, Glenn
; APPLICANT: Sako, Dianne
; APPLICANT: Chang, Xiao Jia
; APPLICANT: Veldman, Geertuida M.
; APPLICANT: Cumming, Dale
; APPLICANT: Kumar, Ravindra
; APPLICANT: Shaw, Gray
; TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESS: LEGAL AFFAIRS
; STREET: 87 CAMBRIDGEPARK DRIVE
; CITY: CAMBRIDGE
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,254A
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/965,662
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/112,608
; FILING DATE: 26-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10168
; FILING DATE: 22-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/225,398
; FILING DATE: 28-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/316,305
; FILING DATE: 30-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWN, SCOTT A.
; REGISTRATION NUMBER: 32,724
; REFERENCE/DOCKET NUMBER: GI 5213E-PCT
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 498-8224
 TELEFAX: (617) 876-5851
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 402 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-477-254A-2

Query Match 100.0%; Score 2030; DB 2; Length 402;
 Best Local Similarity 100.0%; Pred. No. 1.9e-173;
 Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MPQLLLILLIGPNSIQIMDWDADAEKALGPIIARDRROATEYEDYDFLPEETPP 60
 1 MPQLLLILLIGPNSIQIMDWDADAEKALGPIIARDRROATEYEDYDFLPEETPP 60
 61 EMLRNSTDTPLTGPSTPESTVEPARRSTGLDAGAVTELTTELANNGNLSTDSAA 120
 61 EMLRNSTDTPLTGPSTPESTVEPARRSTGLDAGAVTELTTELANNGNLSTDSAA 120
 121 IOTTOPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTP 180
 121 IOTTOPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTP 180
 181 AOTTAPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTP 240
 181 AOTTAPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTP 240
 241 AOTTPLAAMEALSTESATLSEPTTKRGLFIPSSVSVTHKGI PMAASNLVNYYPV 300
 241 AOTTPLAAMEALSTESATLSEPTTKRGLFIPSSVSVTHKGI PMAASNLVNYYPV 300
 301 APHHSVVKOCLLAILLALVATIFVCTVLAVRLSRKGMYPVNNYSPTMVCISLLP 360
 301 APHHSVVKOCLLAILLALVATIFVCTVLAVRLSRKGMYPVNNYSPTMVCISLLP 360
 361 DGEGPSATANGLSKAKSPGLTPEPRDRREGDDLLHSHFLP 402
 361 DGEGPSATANGLSKAKSPGLTPEPRDRREGDDLLHSHFLP 402

RESULT 2
 US-08-472-576B-2
 ; Sequence 2, Application US/08472576B
 ; Patent No. 5840679
 ; GENERAL INFORMATION:
 ; APPLICANT: Larsen, Glenn
 ; APPLICANT: Sako, Dianne
 ; APPLICANT: Chang, Xiao Jia
 ; APPLICANT: Veldman, Geerttruida M.
 ; APPLICANT: Cumming, Dale
 ; APPLICANT: Kumar, Ravindra
 ; APPLICANT: Shaw, Gray
 ; APPLICANT: Camphausen, Raymond
 ; TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN
 ; NUMBER OF SEQUENCES: 34
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LEGAL AFFAIRS
 ; STREET: 87 CAMBRIDGEPARK DRIVE
 ; CITY: CAMBRIDGE
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02140
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/472,576B
 ; FILING DATE:

CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/965,662
 FILING DATE: 23-OCT-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/112,608
 FILING DATE: 26-AUG-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/10168
 FILING DATE: 22-OCT-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/235,398
 FILING DATE: 28-APR-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/316,305
 FILING DATE: 30-SEP-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: BROWN, SCOTT A.
 REGISTRATION NUMBER: 32,724
 REFERENCE/DOCKET NUMBER: GI 5213E-PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 498-8224
 TELEFAX: (617) 876-5851
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 402 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-472-576B-2

Query Match 100.0%; Score 2030; DB 2; Length 402;
 Best Local Similarity 100.0%; Pred. No. 1.9e-173;
 Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MPQLLLILLIGPNSIQIMDWDADAEKALGPIIARDRROATEYEDYDFLPEETPP 60
 1 MPQLLLILLIGPNSIQIMDWDADAEKALGPIIARDRROATEYEDYDFLPEETPP 60
 61 EMLRNSTDTPLTGPSTPESTVEPARRSTGLDAGAVTELTTELANNGNLSTDSAA 120
 61 EMLRNSTDTPLTGPSTPESTVEPARRSTGLDAGAVTELTTELANNGNLSTDSAA 120
 121 IOTTOPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTP 180
 121 IOTTOPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTP 180
 181 AOTTAPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTP 240
 181 AOTTAPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTPAAEAQTTP 240
 241 AOTTPLAAMEALSTESATLSEPTTKRGLFIPSSVSVTHKGI PMAASNLVNYYPV 300
 241 AOTTPLAAMEALSTESATLSEPTTKRGLFIPSSVSVTHKGI PMAASNLVNYYPV 300
 301 APHHSVVKOCLLAILLALVATIFVCTVLAVRLSRKGMYPVNNYSPTMVCISLLP 360
 301 APHHSVVKOCLLAILLALVATIFVCTVLAVRLSRKGMYPVNNYSPTMVCISLLP 360
 361 DGEGPSATANGLSKAKSPGLTPEPRDRREGDDLLHSHFLP 402
 361 DGEGPSATANGLSKAKSPGLTPEPRDRREGDDLLHSHFLP 402

RESULT 3
 US-08-428-734B-2
 ; Sequence 2, Application US/08428734B
 ; Patent No. 5843707
 ; GENERAL INFORMATION:
 ; APPLICANT: Larsen, Glenn
 ; APPLICANT: Sako, Dianne
 ; APPLICANT: Chang, Xiao Jia
 ; APPLICANT: Veldman, Geerttruida M.

APPLICANT: Cumming, Dale
 APPLICANT: Kumar, Ravindra
 APPLICANT: Shaw, Gray
 TITLE OF INVENTION: NOVBL P-SELECTIN LIGAND PROTEIN
 NUMBER OF SEQUENCES: 34
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LEGAL AFFAIRS
 STREET: 87 CAMBRIDGEPARK DRIVE
 CITY: CAMBRIDGE
 STATE: MA
 COUNTRY: USA
 ZIP: 02140
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/428,734B
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/965,662
 FILING DATE: 23-OCT-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/112,608
 FILING DATE: 26-AUG-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/10168
 FILING DATE: 22-OCT-1993
 APPLICATION NUMBER: US 08/235,398
 FILING DATE: 28-APR-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/316,305
 FILING DATE: 30-SEP-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: BROWN, SCOTT A.
 REGISTRATION NUMBER: 32,724
 REFERENCE/DOCKET NUMBER: GI 5213E-PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 498-8224
 TELEFAX: (617) 876-5851
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 402 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-428-734B-2

Query Match 100.0%; Score 2030; DB 2; Length 402;
 Best Local Similarity 100.0%; Pred. No. 1.9e-173;
 Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPEQLLLIILGPGNSLQMDTWADAEKALGPIIARPRORATYEVYLDYDFLPETPEP 60
 DB 1 MPEQLLLIILGPGNSLQMDTWADAEKALGPIIARPRORATYEVYLDYDFLPETPEP 60
 QY 61 EMLRNSTDTPLTGGTPESTTVEPAARRSTGLDAGAVTELTTELANNGNLSTDSAAWE 120
 DB 61 EMLRNSTDTPLTGGTPESTTVEPAARRSTGLDAGAVTELTTELANNGNLSTDSAAWE 120
 QY 121 IOTTGPAATEAQTTPAATEAQTTRILTAATEAQTTPAATEAQTTPAATEAQTTPGTCLE 180
 DB 121 IOTTGPAATEAQTTPAATEAQTTRILTAATEAQTTPAATEAQTTPAATEAQTTPGTCLE 180
 QY 181 AOTTAPAAAEAQTTPAATEAQTTPAATEAQTTPAATEAQTTPAATEAQTTPAATEAQTTPGTCLE 240
 DB 181 AOTTAPAAAEAQTTPAATEAQTTPAATEAQTTPAATEAQTTPAATEAQTTPAATEAQTTPGTCLE 240
 QY 241 AOTTAPAAAEALSTEPSATEALSMEPTTKRGLFIPFSVSSVTHKGI PMASNLVSVY PVG 300

DB 241 AOTTAPAAAEALSTEPSATEALSMEPTTKRGLFIPFSVSSVTHKGI PMASNLVSVY PVG 300
 QY 301 APDHSVKOCILAIILALVATIFVYCVVLAVRLSRKGMV PVRNYSPTMWCISLILP 360
 DB 301 APDHSVKOCILAIILALVATIFVYCVVLAVRLSRKGMV PVRNYSPTMWCISLILP 360
 QY 361 DGGEGPSATANGALSKAKSPGLTPPEBDEBDDTLHSLFLP 402
 DB 361 DGGEGPSATANGALSKAKSPGLTPPEBDEBDDTLHSLFLP 402

RESULT 4
 US-09-063-237-1
 ; Sequence 1, Application US/09063237
 ; Patent No. 6124267
 ; GENERAL INFORMATION:
 ; APPLICANT: McEwer, Rodger P.
 ; TITLE OF INVENTION: O-Glycan Inhibitors of Selectin Mediated
 ; TITLE OF INVENTION: Inflammation Derived from PSGL-1
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Patrea L. Pabst
 ; STREET: 2800 One Atlantic Center, 1201 West Peachtree
 ; CITY: Atlanta
 ; STATE: Georgia
 ; COUNTRY: US
 ; ZIP: 30306-3450
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/063,237
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/649,802
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Pabst, Patrea L.
 REGISTRATION NUMBER: 31,284
 REFERENCE/DOCKET NUMBER: OMRF110CIP7
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (404) 873-8795
 TELEFAX: (404) 873-8795
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 402 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-09-063-237-1

Query Match 100.0%; Score 2030; DB 3; Length 402;
 Best Local Similarity 100.0%; Pred. No. 1.9e-173;
 Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPEQLLLIILGPGNSLQMDTWADAEKALGPIIARPRORATYEVYLDYDFLPETPEP 60
 DB 1 MPEQLLLIILGPGNSLQMDTWADAEKALGPIIARPRORATYEVYLDYDFLPETPEP 60
 QY 61 EMLRNSTDTPLTGGTPESTTVEPAARRSTGLDAGAVTELTTELANNGNLSTDSAAWE 120
 DB 61 EMLRNSTDTPLTGGTPESTTVEPAARRSTGLDAGAVTELTTELANNGNLSTDSAAWE 120
 QY 121 IOTTGPAATEAQTTPAATEAQTTRILTAATEAQTTPAATEAQTTPAATEAQTTPGTCLE 180
 DB 121 IOTTGPAATEAQTTPAATEAQTTRILTAATEAQTTPAATEAQTTPAATEAQTTPGTCLE 180


```

? FILING DATE: 23-OCT-1992
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/112,608
? FILING DATE: 26-AUG-1993
? PRIOR APPLICATION DATA: PCT/US93/10168
? APPLICATION NUMBER: 34
? FILING DATE: 22-OCT-1993
? PRIOR APPLICATION DATA: US 08/235,398
? APPLICATION NUMBER: US 08/235,398
? FILING DATE: 28-APR-1994
? PRIOR APPLICATION DATA: US 08/316,305
? APPLICATION NUMBER: 30-SEP-1994
? FILING DATE: 30-SEP-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: BROWN, SCOTT A.
? REGISTRATION NUMBER: 32,724
? REFERENCE/DOCKET NUMBER: GI 5213B-PCT
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617) 876-5851
? TELEFAX: (617) 876-5851
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 412 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-477-254A-4

```

```

Query Match 99.3%; Score 2015; DB 2; Length 412;
Best Local Similarity 97.6%; Pred. No. 4.3e-172;
Matches 402; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

```

```

QY 1 MPLQLLLLLLILGPNNSLQIMDTWADBAEKALGPIIARDRROATEYEVLDYDFLPETEP 60
DB 1 MPLQLLLLLLILGPNNSLQIMDTWADBAEKALGPIIARDRROATEYEVLDYDFLPETEP 60
QY 61 EMLRNSDTPPLTGTGTPSTVTPVPAARRSTGLDAGAVTELTTELANNGNLSTDSAA 120
DB 61 EMLRNSDTPPLTGTGTPSTVTPVPAARRSTGLDAGAVTELTTELANNGNLSTDSAA 120
QY 121 IOTTPAA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAATE 170
DB 121 IOTTPAA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAATE 170
QY 171 AOTTPGTEAOTTAAMEALSTEPSATEALSMEPTTKRGLFIPPSVSVTHKGI PMAA 230
DB 171 AOTTPGTEAOTTAAMEALSTEPSATEALSMEPTTKRGLFIPPSVSVTHKGI PMAA 230
QY 231 AOTTPGTEAOTTAAMEALSTEPSATEALSMEPTTKRGLFIPPSVSVTHKGI PMAA 290
DB 231 AOTTPGTEAOTTAAMEALSTEPSATEALSMEPTTKRGLFIPPSVSVTHKGI PMAA 290
QY 241 AOTTPGTEAOTTAAMEALSTEPSATEALSMEPTTKRGLFIPPSVSVTHKGI PMAA 300
DB 241 AOTTPGTEAOTTAAMEALSTEPSATEALSMEPTTKRGLFIPPSVSVTHKGI PMAA 300
QY 291 SNLSVNYPVGADPHISVKOCLLAILLALVATTFFVCTVAVLAVRLSRKGMVPRVNYSP 350
DB 291 SNLSVNYPVGADPHISVKOCLLAILLALVATTFFVCTVAVLAVRLSRKGMVPRVNYSP 350
QY 301 SNLSVNYPVGADPHISVKOCLLAILLALVATTFFVCTVAVLAVRLSRKGMVPRVNYSP 360
DB 301 SNLSVNYPVGADPHISVKOCLLAILLALVATTFFVCTVAVLAVRLSRKGMVPRVNYSP 360
QY 351 EMLRNSDTPPLTGTGTPSTVTPVPAARRSTGLDAGAVTELTTELANNGNLSTDSAA 402
DB 351 EMLRNSDTPPLTGTGTPSTVTPVPAARRSTGLDAGAVTELTTELANNGNLSTDSAA 402
QY 361 EMLRNSDTPPLTGTGTPSTVTPVPAARRSTGLDAGAVTELTTELANNGNLSTDSAA 412
DB 361 EMLRNSDTPPLTGTGTPSTVTPVPAARRSTGLDAGAVTELTTELANNGNLSTDSAA 412

```

```

RESULT 7
? US-08-472-576B-4
? Sequence 4, Application US/08472576B
? Patent No. 5840679
? GENERAL INFORMATION:
? APPLICANT: Larsen, Giem
? APPLICANT: Sako, Dianne
? APPLICANT: Chang, Xiao Jia
? APPLICANT: Veidman, Geetruida M.
? APPLICANT: Cumming, Dale
? APPLICANT: Kumar, Ravindra
? APPLICANT: Shaw, Gray

```

```

? APPLICANT: Camphausen, Raymond
? TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN
? NUMBER OF SEQUENCES: 34
? CORRESPONDENCE ADDRESS:
? ADDRESS: LEGAL AFFAIRS
? STREET: 87 CAMBRIDGE PARK DRIVE
? CITY: CAMBRIDGE
? STATE: MA
? COUNTRY: USA
? ZIP: 02140
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/472,576B
? FILING DATE:
? CLASSIFICATION: 435
? APPLICATION NUMBER: US 07/965,662
? FILING DATE: 23-OCT-1992
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/112,608
? FILING DATE: 26-AUG-1993
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PCT/US93/10168
? FILING DATE: 22-OCT-1993
? APPLICATION NUMBER: US 08/316,305
? FILING DATE: 30-SEP-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/235,398
? FILING DATE: 28-APR-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 32,724
? REFERENCE/DOCKET NUMBER: GI 5213B-PCT
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617) 498-8224
? TELEFAX: (617) 876-5851
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 412 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-472-576B-4

```

```

Query Match 99.3%; Score 2015; DB 2; Length 412;
Best Local Similarity 97.6%; Pred. No. 4.3e-172;
Matches 402; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

```

```

QY 1 MPLQLLLLLLILGPNNSLQIMDTWADBAEKALGPIIARDRROATEYEVLDYDFLPETEP 60
DB 1 MPLQLLLLLLILGPNNSLQIMDTWADBAEKALGPIIARDRROATEYEVLDYDFLPETEP 60
QY 61 EMLRNSDTPPLTGTGTPSTVTPVPAARRSTGLDAGAVTELTTELANNGNLSTDSAA 120
DB 61 EMLRNSDTPPLTGTGTPSTVTPVPAARRSTGLDAGAVTELTTELANNGNLSTDSAA 120
QY 121 IOTTPAA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAATE 170
DB 121 IOTTPAA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAATE 170
QY 171 AOTTPGTEAOTTAAMEALSTEPSATEALSMEPTTKRGLFIPPSVSVTHKGI PMAA 230
DB 171 AOTTPGTEAOTTAAMEALSTEPSATEALSMEPTTKRGLFIPPSVSVTHKGI PMAA 230
QY 231 AOTTPGTEAOTTAAMEALSTEPSATEALSMEPTTKRGLFIPPSVSVTHKGI PMAA 290
DB 231 AOTTPGTEAOTTAAMEALSTEPSATEALSMEPTTKRGLFIPPSVSVTHKGI PMAA 290
QY 241 AOTTPGTEAOTTAAMEALSTEPSATEALSMEPTTKRGLFIPPSVSVTHKGI PMAA 300
DB 241 AOTTPGTEAOTTAAMEALSTEPSATEALSMEPTTKRGLFIPPSVSVTHKGI PMAA 300

```

QY 291 SNLSVNVYPVGPADPHISVKQCLLAILLALVATIFVCTVVLAVRLSRKGMYPVRNYSPT 350
 Db 301 SNLSVNVYPVGPADPHISVKQCLLAILLALVATIFVCTVVLAVRLSRKGMYPVRNYSPT 360
 QY 351 EWCISLPLPDGEGSATSATANGLSKXKSPGLTPPEPREDGDDLTLSFLP 402
 Db 361 EWCISLPLPDGEGSATSATANGLSKXKSPGLTPPEPREDGDDLTLSFLP 412

RESULT 8

US-08-428-734B-4
 ; Sequence 4, Application US/08428734B
 ; Patent No. 5843707
 ; GENERAL INFORMATION:
 ; APPLICANT: Larsen, Glenn
 ; APPLICANT: Sako, Dianne
 ; APPLICANT: Chang, Xiao Jia
 ; APPLICANT: Veldman, Geertuida M.
 ; APPLICANT: Cumming, Dale
 ; APPLICANT: Kumar, Ravindra
 ; APPLICANT: Shaw, Gray
 ; APPLICANT: Camphausen, Raymond
 ; TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN
 ; NUMBER OF SEQUENCES: 34
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LEGAL AFFAIRS
 ; STREET: 87 CAMBRIDGEPARK DRIVE
 ; CITY: CAMBRIDGE
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02140
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/428,734B
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/965,662
 ; FILING DATE: 23-OCT-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/112,608
 ; FILING DATE: 26-AUG-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US93/10168
 ; FILING DATE: 22-OCT-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/235,398
 ; FILING DATE: 28-APR-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/316,305
 ; FILING DATE: 30-SEP-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BROWN, SCOTT A.
 ; REGISTRATION NUMBER: 32,724
 ; REFERENCE/DOCKET NUMBER: GI 5213E-PCT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 498-8824
 ; TELEFAX: (617) 876-5851
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 412 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-428-734B-4

Query Match 99.3%; Score 2015; DB 2; Length 412;
 Best Local Similarity 97.6%; Pred. No. 4.3e-172;
 Matches 402; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 MPEQLLILILIGPNSLQIMDTWADAEKALGPLIARDRQATEYEVLDYDFLPETEP 60
 Db 1 MPEQLLILILIGPNSLQIMDTWADAEKALGPLIARDRQATEYEVLDYDFLPETEP 60
 QY 61 EMIRNSTDTPPLTGRPTPESTVVEPARRSTGLDAGGAVTELTTELANKGNLSTDSAME 120
 Db 61 EMIRNSTDTPPLTGRPTPESTVVEPARRSTGLDAGGAVTELTTELANKGNLSTDSAME 120
 QY 121 IOTTOPAA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAATE 170
 Db 121 IOTTOPAA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAATE 180
 QY 171 AOTTOPTEGAEAOITPAAMEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAATE 230
 Db 181 AOTTOPTEGAEAOITPAAMEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAATE 240
 QY 231 AOTTOPTEAOTTPAAMEAALSTEPSATEALSMEPTTRKGLPIPPSVSSVTHKGI 290
 Db 241 AOTTOPTEAOTTPAAMEAALSTEPSATEALSMEPTTRKGLPIPPSVSSVTHKGI 300
 QY 291 SNLSVNVYPVGPADPHISVKQCLLAILLALVATIFVCTVVLAVRLSRKGMYPVRNYSPT 350
 Db 301 SNLSVNVYPVGPADPHISVKQCLLAILLALVATIFVCTVVLAVRLSRKGMYPVRNYSPT 360
 QY 351 EWCISLPLPDGEGSATSATANGLSKXKSPGLTPPEPREDGDDLTLSFLP 402
 Db 361 EWCISLPLPDGEGSATSATANGLSKXKSPGLTPPEPREDGDDLTLSFLP 412

RESULT 9

US-08-713-556F-4
 ; Sequence 4, Application US/08713556F
 ; Patent No. 6277975
 ; GENERAL INFORMATION:
 ; APPLICANT: Larsen, Glenn
 ; APPLICANT: Sako, Dianne
 ; APPLICANT: Chang, Xiao Jia
 ; APPLICANT: Veldman, Geertuida M.
 ; APPLICANT: Cumming, Dale
 ; APPLICANT: Kumar, Ravindra
 ; APPLICANT: Shaw, Gray
 ; TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN
 ; NUMBER OF SEQUENCES: 45
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LEGAL AFFAIRS
 ; STREET: 87 CAMBRIDGEPARK DRIVE
 ; CITY: CAMBRIDGE
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02140
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/713,556F
 ; FILING DATE:
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/965,662
 ; FILING DATE: 23-OCT-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/112,608
 ; FILING DATE: 26-AUG-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US93/10168
 ; FILING DATE: 22-OCT-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/235,398
 ; FILING DATE: 28-APR-1994
 ; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/316,305
 FILING DATE: 30-SEP-1994
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/428,734
 FILING DATE: 25-APR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: BROWN, SCOTT A.
 REGISTRATION NUMBER: 32,724
 REFERENCE/DOCKET NUMBER: GI 5213F
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 498-8224
 TELEFAX: (617) 876-5851
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 412 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-713-556F-4

Query Match 99.3%; Score 2015; DB 4; Length 412;
 Best Local Similarity 97.6%; Pred. No. 4,3e-172;
 Matches 402; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 MPEQLLLILLIGPNSIQLMDTWADAEKALGPIIARDRQATEYEVLDYDFLPETPEPP 60
 |||
 DB 1 MPEQLLLILLIGPNSIQLMDTWADAEKALGPIIARDRQATEYEVLDYDFLPETPEPP 60
 |||
 QY 61 EMLRNSTDTTPLTGPSTPESTTVEPARRSTGIDAGAVTELTTELANGNSTDSAAHE 120
 |||
 DB 61 EMLRNSTDTTPLTGPSTPESTTVEPARRSTGIDAGAVTELTTELANGNSTDSAAHE 120
 |||
 QY 121 IOTTOPA-----TEAOTTPAATEAOTTRLTATEAOTTPAATEAOTTPAATE 170
 |||
 DB 121 IOTTOPA-----TEAOTTPAATEAOTTRLTATEAOTTPAATEAOTTPAATE 180
 |||
 QY 171 AQTTPGTEAQTTPAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPAPEATE 230
 |||
 DB 181 AQTTPGTEAQTTPAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPAPEATE 240
 |||
 QY 231 AQTTPATEAOTTPAAMEALSTEPSATELSMERTTRGPIFIPSSVSYTHKGIPIAMA 290
 |||
 DB 241 AQTTPATEAOTTPAAMEALSTEPSATELSMERTTRGPIFIPSSVSYTHKGIPIAMA 300
 |||
 QY 291 SNLSVNYPVGAPDHISVKOCLLAIIILVALVATIFPVCTVWLAVRLSRKGMYPVRYNSPT 350
 |||
 DB 301 SNLSVNYPVGAPDHISVKOCLLAIIILVALVATIFPVCTVWLAVRLSRKGMYPVRYNSPT 360
 |||
 QY 351 EMVCISSLIPDGGEGSPATANGSLSKAKSPGLTPPEPREDREGDDLTLSHFLP 402
 |||
 DB 361 EMVCISSLIPDGGEGSPATANGSLSKAKSPGLTPPEPREDREGDDLTLSHFLP 412
 |||

RESULT 10
 US-08-713-556F-42
 ; Sequence 42, Application US/08713556F
 ; Patent No. 6277975
 ; GENERAL INFORMATION:
 ; APPLICANT: Larsen, Glenn
 ; APPLICANT: Sako, Dianne
 ; APPLICANT: Chang, Xiao Jia
 ; APPLICANT: Veidman, Geertuida M.
 ; APPLICANT: Cumming, Dale
 ; APPLICANT: Kumar, Ravindra
 ; APPLICANT: Shaw, Gray
 ; TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN
 ; NUMBER OF SEQUENCES: 45
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LEGAL AFFAIRS
 ; STREET: 87 CAMBRIDGEPARK DRIVE
 ; CITY: CAMBRIDGE
 ; STATE: MA
 ; COUNTRY: USA

QY 1 MPEQLLLILLIGPNSIQLMDTWADAEKALGPIIARDRQATEYEVLDYDFLPETPEPP 60
 |||
 DB 1 MPEQLLLILLIGPNSIQLMDTWADAEKALGPIIARDRQATEYEVLDYDFLPETPEPP 60
 |||
 QY 61 EMLRNSTDTTPLTGPSTPESTTVEPARRSTG 92
 |||
 DB 61 EMLRNSTDTTPLTGPSTPESTTVEPARRSTG 92
 |||

RESULT 11
 US-08-713-556F-36
 ; Sequence 36, Application US/08713556F
 ; Patent No. 6277975
 ; GENERAL INFORMATION:
 ; APPLICANT: Larsen, Glenn
 ; APPLICANT: Sako, Dianne
 ; APPLICANT: Chang, Xiao Jia
 ; APPLICANT: Veidman, Geertuida M.
 ; APPLICANT: Cumming, Dale
 ; APPLICANT: Kumar, Ravindra
 ; APPLICANT: Shaw, Gray
 ; TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN
 ; NUMBER OF SEQUENCES: 45
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LEGAL AFFAIRS

QY 1 MPEQLLLILLIGPNSIQLMDTWADAEKALGPIIARDRQATEYEVLDYDFLPETPEPP 60
 |||
 DB 1 MPEQLLLILLIGPNSIQLMDTWADAEKALGPIIARDRQATEYEVLDYDFLPETPEPP 60
 |||
 QY 61 EMLRNSTDTTPLTGPSTPESTTVEPARRSTG 92
 |||
 DB 61 EMLRNSTDTTPLTGPSTPESTTVEPARRSTG 92
 |||

ZIP: 02140
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/713,556F
 FILING DATE:
 CLASSIFICATION: 536
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 07/965,662
 FILING DATE: 23-OCT-1992
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/112,608
 FILING DATE: 26-AUG-1993
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/10168
 FILING DATE: 22-OCT-1993
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/235,398
 FILING DATE: 28-APR-1994
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/316,305
 FILING DATE: 30-SEP-1994
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/428,734
 FILING DATE: 25-APR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: BROWN, SCOTT A.
 REGISTRATION NUMBER: 32,724
 REFERENCE/DOCKET NUMBER: GI 5213F
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 498-8224
 TELEFAX: (617) 876-5851
 INFORMATION FOR SEQ ID NO: 42:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 264 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-713-556F-42

Query Match 22.9%; Score 455; DB 4; Length 264;
 Best Local Similarity 96.7%; Pred. No. 6,9e-34;
 Matches 89; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

? STREET: 87 CAMBRIDGEPARK DRIVE
? CITY: CAMBRIDGE
? STATE: MA
? COUNTRY: USA
? ZIP: 02140
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/713,556F
? FILING DATE:
? CLASSIFICATION: 536
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/965,662
? FILING DATE: 23-OCT-1992
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/112,608
? FILING DATE: 26-AUG-1993
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PCT/US93/10168
? FILING DATE: 22-OCT-1993
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/235,398
? FILING DATE: 28-APR-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/316,305
? FILING DATE: 30-SEP-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/428,734
? FILING DATE: 25-APR-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: BROWN, SCOTT A.
? REGISTRATION NUMBER: 32,724
? REFERENCE/DOCKET NUMBER: GI 5213F
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617) 498-8224
? TELEFAX: (617) 876-5851
? INFORMATION FOR SEQ ID NO: 36:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 313 amino acids
? TYPE: amino acid
? STRANDEDNESS:
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-713-556F-36

```

```

Query Match 22.9%; Score 465; DB 4; Length 313;
Best Local Similarity 97.8%; Pred. No. 8,7e-34;
Matches 89; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MPQLLLLLLILGPGNSLQMDTWADAEKALGPLLARDRQATEYEVLDYDFLPETEP 60
DB 1 MPQLLLLLLILGPGNSLQMDTWADAEKALGPLLARDRQATEYEVLDYDFLPETEP 60
QY 61 EMLRNSTDTTPTLTGPGTPESTTVEPPAR 91
DB 61 EMLRNSTDTTPTLTGPGTPESTTVEPPAR 91

```

```

RESULT 12
US-08-713-556F-38
? Sequence 38, Application US/08713556F
? Patent No. 6277975
? GENERAL INFORMATION:
? APPLICANT: Larsen, Glenn
? APPLICANT: Sako, Dianne
? APPLICANT: Chang, Xiao Jia
? APPLICANT: Veldman, Geertuida M.
? APPLICANT: Cumming, Dale
? APPLICANT: Kumar, Ravindra
? APPLICANT: Shaw, Gray

```

```

? TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN
? NUMBER OF SEQUENCES: 45
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: LEGAL AFFAIRS
? STREET: 87 CAMBRIDGEPARK DRIVE
? CITY: CAMBRIDGE
? STATE: MA
? COUNTRY: USA
? ZIP: 02140
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/713,556F
? FILING DATE:
? CLASSIFICATION: 536
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/965,662
? FILING DATE: 23-OCT-1992
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/112,608
? FILING DATE: 26-AUG-1993
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PCT/US93/10168
? FILING DATE: 22-OCT-1993
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/235,398
? FILING DATE: 28-APR-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/316,305
? FILING DATE: 30-SEP-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/428,734
? FILING DATE: 25-APR-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: BROWN, SCOTT A.
? REGISTRATION NUMBER: 32,724
? REFERENCE/DOCKET NUMBER: GI 5213F
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617) 498-8224
? TELEFAX: (617) 876-5851
? INFORMATION FOR SEQ ID NO: 38:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 269 amino acids
? TYPE: amino acid
? STRANDEDNESS:
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-713-556F-38

```

```

Query Match 22.8%; Score 463; DB 4; Length 269;
Best Local Similarity 100.0%; Pred. No. 1.1e-33;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPQLLLLLLILGPGNSLQMDTWADAEKALGPLLARDRQATEYEVLDYDFLPETEP 60
DB 1 MPQLLLLLLILGPGNSLQMDTWADAEKALGPLLARDRQATEYEVLDYDFLPETEP 60
QY 61 EMLRNSTDTTPTLTGPGTPESTTVEPPAR 88
DB 61 EMLRNSTDTTPTLTGPGTPESTTVEPPAR 88

```

```

RESULT 13
US-08-713-556F-40
? Sequence 40, Application US/08713556F
? Patent No. 6277975
? GENERAL INFORMATION:
? APPLICANT: Larsen, Glenn
? APPLICANT: Sako, Dianne
? APPLICANT: Chang, Xiao Jia

```

APPLICANT: Veldman, Geertuida M.
APPLICANT: Cumming, Dale
APPLICANT: Kumar, Ravindra
APPLICANT: Shaw, Gray
TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN
NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:
ADDRESSER: LEGAL AFFAIRS
STREET: 87 CAMBRIDGEPARK DRIVE
CITY: CAMBRIDGE
STATE: MA
COUNTRY: USA
ZIP: 02140

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/713,556P
FILING DATE:

CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/965,662
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/112,608
FILING DATE: 26-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10168
FILING DATE: 22-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235,398
FILING DATE: 28-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,305
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/428,734
FILING DATE: 25-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: BROWN, SCOTT A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: GI 5213P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
STANDARDS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-713-556F-40

Query Match 13.7%; Score 278; DB 4; Length 437;
Best Local Similarity 92.9%; Pred. No. 7, 1e-17;
Matches 52; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 34 PIIARRRQATEVEYLDYDFLPETPEPMIRNSTDTTPTLGGPTPESTVVEPAARR 89
DB 275 PLIKRERKQATEVEYLDYDFLPETPEPMIRNSTDTTPTLGGPTPESTVVEPAARR 330
PCT-US96-10043-3
Sequence 3, Application PC/TUS9610043
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: P-SELECTIN LIGANDS AND RELATED MOLECULES
NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/10043
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,213
FILING DATE: 14-JUN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Iech, Karen F.
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 00786/284001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
STANDARDS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-10043-3

Query Match 11.4%; Score 232; DB 5; Length 42;
Best Local Similarity 100.0%; Pred. No. 3, 3e-14;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 QIMDIWADBAERKALGPLLARRRQATEVEYLDYDFLPETPEP 60
DB 1 QIMDIWADBAERKALGPLLARRRQATEVEYLDYDFLPETPEP 42
PCT-US96-10043-13
Sequence 13, Application PC/TUS9610043
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: P-SELECTIN LIGANDS AND RELATED MOLECULES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/10043
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,213
FILING DATE: 14-JUN-1995

```

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lech, Karen F.
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 00786/284001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-10043-13

```

```

Query Match 10.6%; Score 216; DB 5; Length 42;
Best Local Similarity 100.0%; Pred. No. 8.9e-13;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Caps 0;

```

```

Qy 60 PEMLRNSTDTTPTLTGPGTPESTTVEPARARSTGIDAGAVTE 101
Db 1 PEMLRNSTDTTPTLTGPGTPESTTVEPARARSTGIDAGAVTE 42

```

```

Search completed: July 14, 2003, 06:16:03
Job time : 46 secs

```